Higgs Production in Association with Top Quarks at the LHC

A thesis submitted to the University of Glasgow in fulfilment of the requirements for the award of the degree of Doctor of Philosophy in the College of Science and Engineering

Gavin Kirby

School of Physics and Astronomy

University of Glasgow

June 2016

Abstract

Since it has been found that the MadGraph Monte Carlo generator offers superior flavour-matching capability as compared to Alpgen, the suitability of MadGraph for the generation of $t\bar{t}b\bar{b}$ events is explored, with a view to simulating this background in searches for the Standard Model Higgs production and decay process $t\bar{t}H, H \to b\bar{b}$. Comparisons are performed between the output of MadGraph and that of Alpgen, showing that satisfactory agreement in their predictions can be obtained with the appropriate generator settings.

A search for the Standard Model Higgs boson, produced in association with the top quark and decaying into a $b\bar{b}$ pair, using 20.3 fb^{-1} of 8 TeV collision data collected in 2012 by the ATLAS experiment at CERN's Large Hadron Collider, is presented. The GlaNtp analysis framework, together with the RooFit package and associated software, are used to obtain an expected 95% confidence-level limit of $4.2^{+4.1}_{-2.0}$ times the Standard Model expectation, and the corresponding observed limit is found to be 5.9; this is within experimental uncertainty of the published result of the analysis performed by the ATLAS collaboration.

A search for a heavy charged Higgs boson of mass $m_{H^\pm}$ in the range $200 \leq m_{H^\pm}/\text{GeV} \leq 600$, where the Higgs mediates the five-flavour beyond-the-Standard-Model physics process $gb \to tH^\pm \to ttb$, with one top quark decaying leptonically and the other decaying hadronically, is presented, using the 20.3 fb^{-1} 8 TeV ATLAS data set. Upper limits on the product of the production cross-section and the branching ratio of the $H^\pm$ boson are computed for six mass points, and these are found to be compatible within experimental uncertainty with those obtained by the corresponding published ATLAS analysis.

Declaration

I declare that, except where otherwise stated, this thesis presents work completed by me, Gavin Kirby, between 2010 and 2015 in the Experimental Particle Physics research group in the School of Physics and Astronomy at the University of Glasgow. None of the work described herein has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

The Monte Carlo validation and physics analysis work presented in Chapters 5, 7, and 8, except where otherwise stated, is my own – the analysis work downstream of the preparation of the ATLAS data and MC files (as used by the ATLAS $t\bar{t}H, H \to b\bar{b}$ working group) was performed by me. The GlaNtp data analysis software framework (as used in the Higgs searches) was developed primarily by Rick St. Denis.

— **Gavin Kirby, 2016-06-17**

Copyright

Colophon

This thesis was typeset by the author using LaTeX 2_ε. Stylistic ideas were appropriated liberally from Andy Buckley's excellent "hepthesis" LaTeX class, which unfortunately the author became aware of too late to implement directly.

Acknowledgements

I am extremely grateful to the Science and Technology Facilities Council and the University of Glasgow for giving me the opportunity to study for this degree.

I would like to acknowledge, in no particular order, the assistance and moral support provided by several people during the course of my PhD studies:

My second supervisor, Tony Doyle, for patiently reading through uncountable thesis drafts and offering constructive feedback.

My examiners, Lars Eklund and David Colling, for their constructive criticisms, which helped me to turn a late draft into a finished thesis.

Andrea Knue, for helping me with analysis and plotting software.

Deepak Kar, for assisting me with my work on Monte Carlo generators.

The members of the ATLAS $t\bar{t}H, H \to b\bar{b}$ analysis group, for our productive collaboration during my sojourn at CERN.

Andrew Pickford, David Crooks, and Gareth Roy, for their help with all manner of computing-related difficulties.

Donny and Alistair, for helping me to troubleshoot GlaNtp.

Will, for his frequent and patient indulgence of my questions, and for sharing his unique humour. The Cult of MOSNUM will rise yet.

Paul, for stimulating office conversation and coding ideas.

SteveØ, for his character.

CCT, for his ever-stimulating chat and decidedly out-of-the-box thinking. Never has the discovery of a USB flash drive provided so much entertainment.

The CERN LTA folks during my era (2011–2013), for all the laughs and the good times.

Maciej, for entertaining me while I was finishing up.

Ruaridh, for sharing my sense of humour.

J&S, for being family.

And anyone who seriously entertained my pub ramblings at any time; it was all only "a jest in sober earnest".

"All reality is a game. Physics at its most fundamental, the very fabric of our universe, results directly from the interaction of certain fairly simple rules, and chance; the same description may be applied to the best, most elegant and both intellectually and aesthetically satisfying games. By being unknowable, by resulting from events which, at the sub-atomic level, cannot be fully predicted, the future remains malleable, and retains the possibility of change, the hope of coming to prevail; victory, to use an unfashionable word. In this, the future is a game; time is one of its rules."
— Iain M. Banks, *The Player of Games*

"Ich sage euch: man muß noch Chaos in sich haben, um einen tanzenden Stern gebären zu können."
— Nietzsche, *Also Sprach Zarathustra*

This thesis is dedicated to all those who helped and encouraged me during my time at Glasgow and at CERN.

To the memory of my original supervisor, Rick St. Denis.

And most of all to my loving parents, without whose unfailing support its completion would have been impossible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> "A beginning is the time for taking the most delicate care that the balances are correct."
>
> — Frank Herbert, *Dune*

During its first run from 2009 to 2013 [1], the Large Hadron Collider (LHC) [2, 3] at the European Organisation for Nuclear Research (CERN) produced unprecedented amounts of high-energy proton-proton collision data. Its physics objectives are extremely wide-ranging and ambitious – it aims to allow the exploration of physics at the teraelectronvolt (TeV) scale, with the potential for observing new physics at these hitherto-inaccessible energies. The most notable event of this period was the announcement by the ATLAS [4] and CMS [5] experiments in July 2012 of the observation of a new neutral scalar boson consistent with the Standard Model (SM) Higgs boson [6, 7] – a previously unobserved particle, the existence of which had been predicted by theory since the 1960s. This particle is an important part of the Standard Model of particle physics [8] – the basis of current understanding of physics at the smallest scales – due to its role in explaining the spontaneous breaking of electroweak symmetry (hence accounting for the differing masses of the electroweak gauge bosons $W^{\pm}$, Z, and γ[(i)]) and it was also the last remaining undiscovered fundamental particle according to the Standard Model.

This announcement represented the achievement of one of the LHC's main objectives, and the natural continuation of this line of research was to study the properties of this new boson in greater detail; for example, the Higgs self-coupling remains to be measured. Moreover, not all theoretically allowed modes of Higgs prodution have been observed thus far, and in particular it remains to be understood how strongly

[(i)] This thesis follows the convention of denoting particles with italic symbols.

the newly observed boson couples to other particles – especially the fermions such as the heaviest Standard Model particle, the top quark. Since the Standard Model Higgs couples to other particles according to their masses, it is anticipated that the coupling between the Higgs and the top quark will provide the readiest source of information about these couplings.

This thesis describes work undertaken to explore the viability of some potential discovery modes for Higgs production in association with the top quark, considering both the Standard Model Higgs scenario with the observed mass[ii] $m_H = 125\,\text{GeV}$ and the $H \to b\bar{b}$ decay mode, as well as a range of theoretical scenarios involving charged Higgs bosons with different masses: these multivariate analyses are performed by means of a neural network. An overview of the properties and capabilities of the LHC accelerator and the ATLAS detector are given, together with a description of the theoretical underpinnings of the Standard Model, with a particular focus on electroweak and Higgs physics. The statistical methods and the analysis software employed in the presented analyses are outlined, and computational methods underlying the production of simulated data (Monte Carlo) for use in ATLAS analyses are described.

[ii] This thesis uses the system of natural or Planck units, in which $\hbar = 1$ and $c = 1$ by definition, so "GeV" can be read as equivalent to "GeV/c" in the case of momentum and "GeV/c^2" in the case of mass. Except where otherwise specified, quantities such as p_T and m_H have units of GeV. The signature of the metric tensor is taken to be $(-,+,+,+)$.

Chapter 2

The Standard Model and Higgs Searches

"This summer I have discovered something totally useless."

— Peter Higgs, 1964

"Three quarks for Muster Mark!"

— James Joyce, *Finnegans Wake*

2.1 The Standard Model

2.1.1 Introduction – bosons and the fundamental forces

The Standard Model of particle physics is a theory describing interactions between fundamental particles of half-integer spin (the fermions) in terms of fundamental particles of integer spin (the bosons), which are understood to mediate the fundamental forces. The Standard Model has been extremely successful in describing a multitude of decay and scattering processes at all energy regimes that have been experimentally probed thus far, hence allowing the identification of three fundamental physical interactions (*i.e.* neglecting gravity, which is not described by the Standard Model) which are mediated by gauge bosons. These are the familiar electromagnetic (EM) interaction (mediated by the photon, γ), the weak interaction (mediated by the W^+, W^-, and Z bosons) and the strong interaction (mediated by the gluon, g) [8]. These bosons are detailed in Table 2.1, along with the hypothetical graviton (G) – a particle outside the Standard Model which is postulated to mediate gravitational interactions [9].

Force	Boson	Charge	Spin	Mass (GeV)
Strong	gluon (g)	0	1	0
EM	photon (γ)	0	1	0
Weak	$W^{\pm}, Z$	$\pm 1, 0$	1	$80, 91$
Gravity	graviton (G)	0	2	0

Table 2.1: A table showing the bosons responsible for mediating the four fundamental forces, with some of their main properties. For the gluon and the electroweak gauge bosons, these properties have been experimentally validated; for the graviton they are based on theoretical expectation.

Electromagnetic interactions are responsible for all of chemistry, and hence (neglecting gravity) they account for all known behaviour of matter at electron-volt energy scales (equivalently, at distances greater than the dimensions of the atom). The electromagnetic interaction is mediated by the photon, a massless vector boson; because the photon is massless and lacks self-interactions, the electromagnetic potential scales with distance r as $\frac{1}{r}$.

The weak force accounts for the decay of massive particles such as the heavier baryons and the top quark (which almost invariably decays via the mode $t \to bW^+$) [10]. It also plays an important role in fusion processes in star cores. Because it is mediated by massive vector bosons, its associated potential scales as $\frac{e^{-mr}}{r}$ for force-carrying bosons of mass m – *i.e.* an exponential decay with distance, resulting in an interaction with an extremely short range.

The strong force binds colour-charged objects (quarks and gluons, collectively *partons*) together in composite states known as hadrons. Two broad families of hadrons exist in the Standard Model, characterised by their numbers of valence quarks (*i.e.* the quarks within the hadron which determine their quantum numbers – this excludes so-called *virtual* or *sea* quarks, of which a hadron may contain an indefinite number): states with two valence quarks are bosons and are termed *mesons*, whereas states with three valence quarks are fermions and are termed *baryons*. This latter classification includes the most familiar, stable and important hadrons, the proton (uud) and the neutron (ddu).

In the Standard Model, neutrons undergo weak decay (*i.e.* β decay) to protons over a timescale of minutes, whereas protons are unable to decay to any other state as they are the least massive baryon, and since baryon number is a conserved quantity in the Standard Model, decays to mesons are impermissible. However, in some

beyond-the-Standard-Model (BSM) theories, it is possible for protons to decay (*e.g.* $p \to l^+\pi^0$), which makes limit-setting on the proton lifetime important for validating the Standard Model [11]. The Standard Model neither implies nor precludes the existence of so-called "exotic" hadrons, *i.e.* bound states with more than three valence quarks, such as *tetraquarks* of four valence quarks and *pentaquarks* of five – indeed there is experimental evidence for the existence of both of these types of exotic state [12–14].

The strong force has some unique properties. The boson associated with the strong force, the gluon, is the only force carrier with a self-coupling, because gluons are themselves colour-charged[(i)]. As a consequence, the potential associated to the strong interaction grows with distance due to the "anti-screening" effect of the cloud of colour charge from virtual gg pairs surrounding a bare parton, which serves to enhance the perceived strength of the colour-charged object. Thus, the effective strong coupling constant α_s appears greater at larger distances (equivalently, $\alpha_s\left(Q^2\right)$ decreases at higher energy transfer Q^2 – the so-called asymptotic freedom of the running strong coupling). This very property, however, results in a counter-intuitive limiting of the strong force's effective range, because when two strongly-interacting objects are separated by more than a certain small distance, the increase in strong potential corresponding to the increased separation will be sufficient to create more colour-charged objects from the vacuum, resulting in hadronisation – *i.e.* the formation of colourless states (or *colour singlets*) from the colour-charged fundamental particles, quarks and gluons. Because of this phenomenon it is impossible to observe colour-charged states, such as bare quarks and gluons, directly; this is termed *colour confinement*. The experimental observation corresponding to this hadronisation process is the *jet*: the cone-like shower of hadrons coming radially outwards from the interaction point. By analysing the substructures of these jets it is possible to construct a probability function to describe how likely it is that a given jet originated from a particular kind of hadron; this approach is particularly important in identifying jets that result from the hadronisation of b-quarks (*i.e.* b-tagging).

The remaining fundamental force is gravity, which is often described as being mediated by a spin-2 boson called the graviton, since this is the most natural quantum field theory interpretation of the Einstein field equations [9]. Notably, the graviton does not form part of the Standard Model, and there is currently no direct evidence for its existence; owing to the extreme weakness of gravity compared to

[(i)] There are three types of colour charge – red, green and blue. Gluons carry one colour charge and one anti-colour charge; quarks have a single charge of either kind, depending on whether they are antiquarks.

the other three interactions at the length scales typically probed in particle physics experiments, the effects of gravity are very difficult to measure.

With the exception of the hypothetical graviton, all force-carrying bosons are vector bosons, meaning that they have a spin of one. The Higgs, as will be described later, is different; it is a spinless (or *scalar*) boson.

2.1.2 Fermions – Quarks and Leptons

Whereas the bosons mediate the interactions between matter particles (and between the bosons themselves), matter is constituted by a different set of fundamental particles termed fermions. These are particles of half-integer spin, which therefore behave quite differently from bosons. The chief difference in relativistic quantum field theory is that fermionic states ψ are thought of as solutions to the Dirac equation

$$\left(\mathrm{i}\not{\partial} - m\right)\psi = \left(\mathrm{i}\partial^{\mu}\gamma_{\mu} - m\right)\psi = 0, \tag{2.1}$$

and they obey Fermi-Dirac statistics, whereas bosons are solutions to the Klein-Gordon equation

$$\left(\Box + \mu^2\right)\psi \equiv \left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\psi = 0, \tag{2.2}$$

(where $\Box \equiv -\eta^{\mu\nu}\partial_{\mu}\partial_{\nu} = \frac{\partial}{\partial t} - \nabla^2$ denotes the d'Alembertian operator, *i.e.* the Laplacian of Minkowski space) and they obey Bose-Einstein statistics [15].

One of the consequences of the bosons' obeying Bose-Einstein statistics is that it is possible for several bosons to occupy any given quantum state, whereas this is forbidden for fermions (the Pauli exclusion principle). A consequence of this latter fact is that fermions, unlike bosons, "occupy" volume in space, and it is this fundamental property which causes matter to be extended rather than point-like.

The fundamental division within the fermions is between the quarks and the leptons. Quarks are fermionic partons (colour-charged objects), whereas leptons lack colour charge and hence interact exclusively via the electromagnetic and weak interactions. Leptons which lack electric charge are termed neutrinos, and these interact solely via the weak interaction. Initially believed to be massless, it has in recent decades been established empirically that neutrino masses are nonzero [16] (due to the "mixing" of mass eigenstates, which implies a mass hierarchy, and which is forbidden if the

neutrino masses are all zero) but too small to be measured directly with presently available experimental techniques [8].

The known fundamental fermions are divided into three generations of matter. For the leptons, these are the electron, the muon, the tau, their respective neutrinos, and the antiparticles of these six particles. For the quarks, each generation features an "up-type" quark with charge $+\frac{2}{3}$ and a "down-type" quark with charge $-\frac{1}{3}$. These six quarks are the up, the down, the charm, the strange, the top and the bottom.

An important property used to classify these particles is their helicity, defined by $h = \vec{S} \cdot \hat{p}$, where $\vec{S}$ is the particle's spin vector and $\hat{p}$ is a unit vector in the direction of its momentum; the sign of the helicity describes whether a state is left-handed (*i.e.* the particle's spin vector opposes the direction of its momentum vector) or right-handed (*i.e.* the spin and momentum vectors are parallel). For massive particles, however, this quantity is not well-defined, since it is always possible to Lorentz-boost into a reference frame in which the particle is propagating in the direction $-\hat{p}$. Thus one can define a more general concept of *chirality* or "handedness" based on whether a particle transforms according to a left- or right-handed representation of the Poincaré group; this formalises the idea of a state with the property that, like a hand, it is intrinsically not identical to its mirror image.

Most particles exist in both left-handed and right-handed chirality states; however, the chiral nature of the electroweak theory implies that only left-handed fermions may interact via the weak interaction. Since neutrinos cannot interact via any other interaction, it has not yet been possible to observe right-handed neutrinos, and they are taken not to exist within the Standard Model. Thus, all fermions are represented as either left-handed doublets or right-handed singlets. The left-handed lepton states are shown in (2.3), and the corresponding right-handed states are shown in (2.4). The left-handed quark states are shown in (2.5) and the right-handed quark states in (2.6) and (2.7). Notably, the left-handed quarks comprise three doublets, whereas the right-handed quarks comprise two singlets.

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_{\mathrm{L}} \tag{2.3}$$

$$\begin{pmatrix} e & \mu & \tau \end{pmatrix}_{\mathrm{R}} \tag{2.4}$$

$$\begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} c \\ s \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} t \\ b \end{pmatrix}_{\mathrm{L}} \tag{2.5}$$

$$\begin{pmatrix} u & c & t \end{pmatrix}_{\mathrm{R}} \tag{2.6}$$

$$\begin{pmatrix} d & s & b \end{pmatrix}_{\mathrm{R}} \tag{2.7}$$

2.2 Symmetry

Any physical field theory can be formulated in terms of Lagrangian mechanics, according to which the theory describes physics in terms of a Lagrange density (or, simply, Lagrangian). Field theories of which the Lagrangians are invariant under a certain group of transformations are referred to as gauge theories; the Standard Model is an example of such a gauge theory. The Lagrangians associated to each of the three fundamental forces within the SM are each invariant under their own set of transformations, corresponding to their respective symmetry groups. Combining these three symmetry groups, the overall symmetry group of the Standard Model Lagrangian is $\mathrm{SU}(3) \times \mathrm{SU}(2)_{\mathrm{L}} \times \mathrm{U}(1)_{\mathrm{Y}}$, comprising the product of the respective symmetry groups of the strong force, the electromagnetic force and the weak force.

2.2.1 $\mathcal{CP}$ Symmetry

Each particle in the Standard Model has a corresponding antiparticle, which has identical properties except for its electric charge, flavour, and lepton numbers, which are exactly opposite. The neutral bosons γ, Z, and g are each their own respective antiparticles (fermions with this property are termed *Majorana* particles, although none have yet been observed); the charged W^+ and W^- bosons are each other's antiparticles. By contrast, almost all fermions are known to be *Dirac* particles, meaning that they have distinct antiparticle states; the neutrinos are the exception, as it is currently unknown whether they are Majorana or Dirac particles.

$\mathcal{CP}$ (charge conjugation – parity) symmetry is the symmetry that is exhibited between these antiparticles and their associated particles; if it were an exact symmetry, then physics would be indistinguishable if all particles in the universe were interchanged with their respective antiparticles, and all spatial dimensions were simultaneously inverted. In reality, matter and antimatter are not interchangeable in this way; clearly, there is much more of the former than the latter in the Universe, and the source of this asymmetry is not well understood.

2.2.2 Spontaneous Electroweak Symmetry Breaking – The Higgs Mechanism

During the development of modern physics the importance of fundamental symmetries of nature became increasingly apparent. These include the well-known Poincaré symmetry group of special relativity, as well as the gauge symmetries of quantum field theories. It is well-established that the electromagnetic and weak interactions combine above a certain energy threshold, forming the electroweak interaction; the electromagnetic interaction itself is a unification of the electrostatic and magnetic forces.

Moreover, it became apparent during the development of the Standard Model that certain fundamental symmetries of nature would need to be broken in order to produce the physics that is observed at lower energies, hence the concept of symmetry breaking. This is needed in the electroweak model because, if electroweak symmetry were unbroken, all the fundamental Standard Model particles would be massless, which is clearly inconsistent with observation: in particular, massive vector bosons are needed to account for the "weakness" and short range of the weak force. This difficulty cannot be circumvented by inserting "bare" mass terms into the electroweak Lagrangian, since doing so would violate gauge invariance and make the theory non-renormalisable; instead, the most parsimonious explanation is that the vector boson masses are generated dynamically by some other physical interaction.

Beginning in the 1960s, theorists[ii] developed field theories in which a unified interaction may "split" into two separate interactions through a process known as spontaneous symmetry breaking [17–19]. In the case of the electroweak interaction, this process of symmetry breaking gives rise (at suitably low energies) to two distinct physical interactions – *i.e.* the electromagnetic and weak interactions. The symmetry group associated to the electroweak interaction changes accordingly, with the original $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ symmetry group (where $Y = 2(Q - T_3)$, with T_3 the third component of weak isospin, is the weak hypercharge) being broken down to the global U(1) symmetry of the electromagnetic interaction. The physical consequences of this can be understood by considering the behaviour of the associated vector bosons. Prior to symmetry breaking, the electroweak interaction has four degrees of freedom, which can be understood as three vector bosons W^+, W^-, and W^0 from $\mathrm{SU(2)_L}$ that couple left handed states (left-handed leptons and neutrinos

[ii] In order to give credit to other responsible theorists, the name of the Higgs mechanism is sometimes expanded to include names such as Brout and Englert. For the sake of completeness, Peter Higgs has proposed referring to it as the "ABEGHHK'tH mechanism".

– there are no right-handed neutrinos in the Standard Model) to one another, and one vector boson B^0 from $\mathrm{U(1)_Y}$ that couples left- and right-handed leptons (*i.e.* states of nonzero Y) to one another. There are two associated couplings: that of the three W fields to the fermions, denoted g_W, and that of the B^0 field to the fermions, denoted g'_W. Since B^0 and W^0 have the same quantum numbers, they mix to give the two physical bosons γ and Z^0 in a manner that can be formulated as a rotation in phase space according to the equations

$$\begin{aligned} \gamma &= \cos\theta_{\mathrm{W}}\, B^0 - \sin\theta_{\mathrm{W}}\, W^0 \\ Z^0 &= \sin\theta_{\mathrm{W}}\, B^0 + \cos\theta_{\mathrm{W}}\, W^0, \end{aligned} \tag{2.8}$$

where θ_{W} denotes the (Weinberg) weak mixing angle, which parametrises the mixing [20].

This symmetry breaking is associated with a certain energy scale – the electroweak scale, $v = \left(\sqrt{2}\, G_{\mathrm{F}}\right)^{-\frac{1}{2}}$, where $G_{\mathrm{F}} \approx 1.17 \times 10^{-5}\,\mathrm{GeV}^{-2}$ denotes the Fermi coupling [21]. Within the Higgs model, this can be interpreted as the vacuum expectation value (denoted "vev" or v) of the Higgs field, and its nonzero value (the height of the peak in the centre of the "Mexican hat" potential as illustrated in Figure 2.1) is what allows for the breaking of the electroweak symmetry. Together with the gauge couplings to the W and B fields, v is one of three "free" (theoretically undetermined) parameters in the electroweak model (of course, other choices of the three parameters are possible – for experimental purposes the most precisely measured quantities are typically used, namely M_Z, e and G_{F}) [8].

Different mechanisms have been postulated to account for the breaking of electroweak symmetry, including various Higgs models and technicolor [22–24]. The Higgs model[(iii)] allows for these masses to be generated dynamically via interactions between the other Standard Model particles and a scalar Higgs field. Within this model, the physical Higgs field is the remnant of a complex scalar SU(2) Higgs doublet which possesses four intrinsic degrees of freedom, three of which mix with the other fields in the electroweak model to give the physical (massive) W and Z bosons, while the fourth component is understood to act as a separate physical field, i.e. the Higgs boson. The Higgs term in the electroweak Lagrangian introduces two additional self-coupling parameters: the quadratic coupling μ and the quartic coupling λ, which is related to the Higgs vev v and the Higgs mass m_H by the relation

[(iii)] When stated without qualification in this thesis, "the Higgs boson/model" refers to the Standard Model Higgs.

$m_H = v\sqrt{2\lambda}$; μ is left as a free parameter and must be measured independently [20].

For a complex scalar Higgs field Φ it is possible to define a Lagrangian $\mathcal{L}$ with (real-valued) kinetic and potential terms,

$$\mathcal{L} = T(\Phi) - V(\Phi) = (\partial_\mu \Phi)(\partial^\mu \Phi)^* - \left(\mu^2 \Phi\Phi^* + \lambda(\Phi\Phi^*)^2\right). \tag{2.9}$$

The shape of the Higgs potential $V(\Phi)$ then depends on the sign of μ^2 (λ must not be negative, in order to ensure that the potential is bounded). In the case of $\mu^2 > 0$, there is a unique vacuum solution $V|_{\Phi=0} = 0$; however, in the case of $\mu^2 < 0$, the "trivial" solution at $\Phi = 0$ is seen to be a "false" vacuum, since although it represents a metastable state of the field, it is not the true minimum of potential but rather a local maximum, dropping down to the true minimum in all directions; this is the so-called "Mexican hat" shape. If we consider the true minimum of potential as the "true" physical vacuum of the Higgs field, then we see that the false vacuum is separated from this true vacuum state by an amount of energy

$$\langle 0|\Phi|0\rangle = \frac{|\mu|}{\sqrt{\lambda}} = v \approx 246\,\mathrm{GeV}, \tag{2.10}$$

the Higgs vacuum expectation value. It is this discrepancy between the false (metastable, associated with a relative maximum of energy) and true (truly stable, globally-minimal energy state of the field) vacua that allows spontaneous symmetry breaking to occur; the imperfect stability of the false vacuum state implies that the Higgs field will inevitably drop down into its true minimum when the system is subjected to any perturbation; as illustrated in Figure 2.1, this can be thought of as analogous to a marble, balanced on top of an upturned bowl, falling towards the "true" zero of gravitational potential on the ground and away from the "false" zero of potential on which it had been balanced. The marble's "choice" of direction is what then breaks the symmetry of that initially-symmetric system.

As in the case of the marble, the Higgs field actually has an infinite number of minimal-energy states; these can be represented as a circle in the complex plane with radius v, such that $\Phi\Phi^* = (\Re(\Phi))^2 + (\Im(\Phi))^2 = v^2$. These minima are specified by

$$\Phi_{\min} = \mathrm{e}^{\mathrm{i}\theta}\sqrt{\frac{\mu^2}{2\lambda}}\,, \tag{2.11}$$

where θ represents the angle around the axis of potential. Clearly, any specific

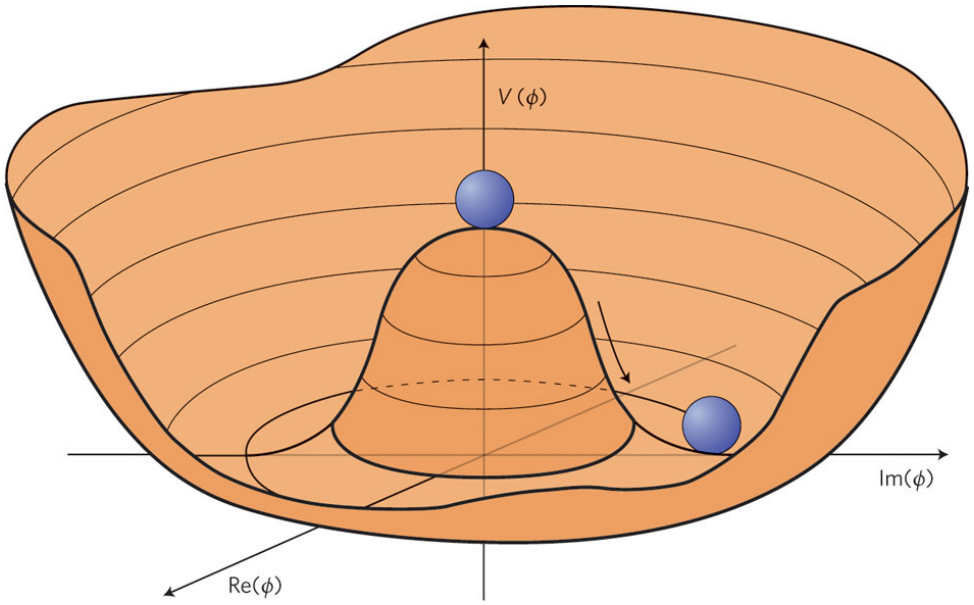


Figure 2.1: An illustration of the "Mexican hat" potential $V(\Phi)$ as a real-valued function in the complex plane $(\Re(\Phi), \Im(\Phi))$ – note that the vacuum expectation value is nonzero, as illustrated by the elevated central region. Taken from Ref. [25].

choice of θ will result in the invariance of the system being lost (specifically, the U(1) invariance). The question then is how to define a gauge for this theory; by convention, the choice is to set $\theta = 0$ so then $\Phi_{\text{min}} = \Phi_{\text{vac}} = \frac{v}{\sqrt{2}}$ is the value of the potential in the vacuum, and the physical vacuum Higgs field is correspondingly reduced from a complex scalar weak isospin doublet to simply

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} , \tag{2.12}$$

i.e. a form in which it possesses only one degree of freedom. Thus, spontaneously breaking the global U(1) symmetry by "choosing" the direction in which to fluctuate has "eaten" three of the four degrees of freedom possessed intrinsically by the Higgs field.

It is now clear that there are many possible excitations of the system which will perturb it away from its vacuum state, and hence towards one of the minimal-energy states. These excitations can be associated with zero energy, *i.e.* associated with massless particles. Such a massless particle would be the Goldstone boson of this theory – a type of scalar particle which always appears in any theory which features a spontaneously broken symmetry. Since the symmetry is not broken explicitly by terms in the Lagrangian of the theory, the boson must be massless.

The field Φ can then be altered such that

$$\Phi = \frac{1}{\sqrt{2}} \left(\frac{\mu}{\sqrt{\lambda}} + H + \mathrm{i}\phi \right), \tag{2.13}$$

where H and the new field ϕ both have zero vacuum expectation value. The potential term in the Lagrangian can then be rewritten as

$$V = \mu^2 H^2 + \mu\sqrt{\lambda} \left(H^3 + \phi^2 H \right) + \frac{\lambda}{4} \left(H^4 + \phi^4 + 2H^2\phi^2 \right) + \frac{\mu^4}{4\lambda}. \tag{2.14}$$

Note that while H has an associated mass term $\mu^2 H^2$, there is no equivalent mass term for the field ϕ.

However, it follows that since these Goldstone bosons would be able to transform the vacuum of the theory into one of the other degenerate minimal-energy states (*i.e.* an unphysical transformation which should be disallowed by the initial choice of gauge specified by $\theta = 0$), they should not be considered as physical particles. Rather, the additional degrees of freedom in electroweak theory that the Goldstone bosons would have "used up" are used instead by the gauge bosons of the electroweak theory.

In total there are four of these degrees of freedom in the electroweak theory (since Φ is actually a doublet of complex scalars, with two real parts and two imaginary parts). Three of these degrees of freedom are used by the W and Z bosons, and the fourth is associated to the massive scalar boson H above, *i.e.* the Higgs. Since the coupling of the photon to the Higgs field is zero, its mass is also zero. One can then deduce these relations between the W and Z boson masses m_W and m_Z, the Higgs mass m_H, the Higgs vev v and the gauge couplings g_W and g'_W:

$$\begin{aligned} m_Z &= v \, \frac{\sqrt{g_W^2 + g_W'^2}}{2}, \\ m_W &= v \, \frac{g_W}{2}, \\ m_H &= v\sqrt{2\lambda} = \sqrt{2}\,|\mu|. \end{aligned} \tag{2.15}$$

This description so far covers only how the masses of the electroweak bosons could be generated dynamically. It is also possible to account for fermion masses m_f by introducing a new set of couplings between the fermions and the Higgs field: these are the Yukawa couplings λ_f. In terms of these couplings, then, the fermion masses can be related to v by

$$m_f = v \frac{\lambda_f}{\sqrt{2}}. \tag{2.16}$$

2.2.3 Supersymmetry

A significant theoretical problem with the Standard Model, including only one light Higgs boson, is that certain quantum field theory terms which contribute to the Higgs mass diverge quadratically at high energy scales, making the observed low-mass Higgs scenario difficult to justify theoretically without resorting to an *ad hoc* fine-tuning whereby the bare-mass terms and the quadratic terms cancel each other to an implausible degree of precision, leaving a residual physical mass which is extremely small relative to the necessary theoretical correction. This, the so-called *hierarchy problem*, is a powerful argument that the Standard Model is merely a low-energy effective theory which cannot provide a satisfactory and consistent model of physics up to the Planck scale.

One major theoretical proposal which represents a potential solution to the hierarchy problem without invoking an extremely precise renormalisation is to consider a new fundamental symmetry of nature: a symmetry between fermions and bosons, termed *supersymmetry* or *SUSY* [26]. According to supersymmetry, each elementary boson has an equivalent fermion superpartner ("bosino") and each elementary fermion has an equivalent boson superparner ("sfermion"). The introduction of these superpartners provides an elegant solution to the hierarchy problem: each of the previously divergent terms in the expansion of the quadratic Higgs mass term acquires an equivalent negative term corresponding to a direct cancellation by its superpartner (as a consequence of the spin-statistics theorem, opposite signs are assigned to terms which represent fermions and bosons respectively). However, as of 2016, no such superpartners have been observed; in order to allow for the non-existence of superpartners at explored energy scales, supersymmetry is considered as a broken symmetry, meaning that the masses of the superpartners must be considerably greater than those of the Standard Model particles, and hence inaccessible at the energy scales that colliders have been able to probe [27]. The difference in mass between the top (the heaviest quark) and the stop (the lightest squark) is key: the stop is theoretically able to stabilise the Higgs vacuum against divergences, provided that the stop mass is close to the electroweak scale (246 GeV). Moreover, the potential relevance of supersymmetry to other theoretical problems such as the composition of dark matter, grand unification, the stability of the proton, and string

State	Description
G^0	neutral Goldstone boson
$G^\pm$	charged Goldstone bosons
h^0	light $\mathcal{CP}$-even Higgs boson
A^0	neutral $\mathcal{CP}$-odd Higgs boson
H^0	heavy neutral $\mathcal{CP}$-even Higgs boson
$H^\pm$	heavy charged $\mathcal{CP}$-even Higgs bosons

Table 2.2: A table showing the eight electroweak degrees of freedom in the MSSM.

theory, makes it of great interest to physicists, and many experimental searches for evidence of supersymmetry are ongoing [28].

2.2.4 The MSSM

The Minimal Supersymmetric Standard Model (MSSM) is the most conservative model to incorporate supersymmetry within the framework of established Standard Model physics; it permits only the bare minimum of new states and interactions required to do so [26]. Table 2.2 describes the eight degrees of electroweak freedom present in the MSSM, deriving from two complex doublets: only the first four are present in the SM, while the latter four correspond to new, more massive Higgs bosons. The lightest Higgs boson h^0 corresponds to the Standard Model Higgs, while the neutral and charged Goldstone bosons correspond to the Z and W bosons respectively.

As with the Standard Model, it is not possible to compute expected masses for the MSSM Higgs bosons directly from theory. Rather, the MSSM Higgs sector is governed by two free parameters: the mass of the charged Higgs bosons $m_{H^\pm}$ and the ratio of the vacuum expectation values of the two Higgs fields arising from the two complex doublets, denoted

$$\tan\beta = \frac{v_2}{v_1} \tag{2.17}$$

where v_1 and v_2 are related to the Standard Model Higgs vacuum expectation value v_{SM} by $v_{\mathrm{SM}}^2 = v_1^2 + v_2^2$.

2.3 Higgs Searches

If the Higgs behaves as the Standard Model predicts, then it should manifest itself in the form of corrections to the cross sections of certain high-energy scattering processes, such as could be studied at a suitably powerful collider. This is how direct searches can be carried out. Indirect searches are also possible, since it is possible for Higgs loops to cause higher-order corrections to the values of electroweak observables; this was an important approach adopted by Higgs searches using LEP and Tevatron data. These precise measurements of the W and top masses, statistically combined with the results of direct searches, were used to constrain the Higgs mass to the range $114\,\text{GeV} \leq m_H \leq 157\,\text{GeV}$ at the 95 % confidence level [29, 30].

In addition, the Higgs model provides an important mechanism for protecting the unitarity of the Standard Model. Because of self-interactions, the cross section of longitudinal WW scattering would be divergent at some energy scale $\Lambda \gtrsim 1\,\text{TeV}$ without some additional corrections at higher orders in perturbation theory to suppress this behaviour [31, 32]. This gives an important theoretical constraint on the Higgs mass, since it can be shown that in order for the Higgs to prevent this divergence, its mass must satisfy the constraint

$$m_H^2 \leq \frac{8\pi^2 v^2}{2\ln\left(\Lambda^2\right)/v^2}\,. \tag{2.18}$$

One further theoretical bound on m_H is the triviality requirement. This means that the Higgs self-coupling λ must not reach a so-called Landau Pole, *i.e.* an energy at which the coupling would be infinite [33]. The Higgs mass can therefore be constrained by imposing an upper limit on the energy at which the SM is valid. There is also the requirement that the electroweak vacuum be stable; the constraint that is derived from this depends on the choice of λ; lower energies will give a weaker constraint.

It is particularly important to consider the branching ratios of the different Higgs decay modes, as shown in Figure 2.2. As an uncharged scalar boson, the Higgs can decay into any particle-antiparticle pair (where the particle has a mass less than half the Higgs mass, in order to conserve energy), but the relative frequencies of these different decay modes vary significantly with m_H. Notably, searches for "light" Higgs ($m_H < 135\,\text{GeV}$) before discovery focussed on the decay modes $H \to \gamma\gamma$ (mediated by a fermion loop), $H \to \tau^+\tau^-$ and $H \to b\bar{b}$ (the difficulty in distinguishing charm jets from other light jets presents a large barrier to the study of $H \to c\bar{c}$ decays for

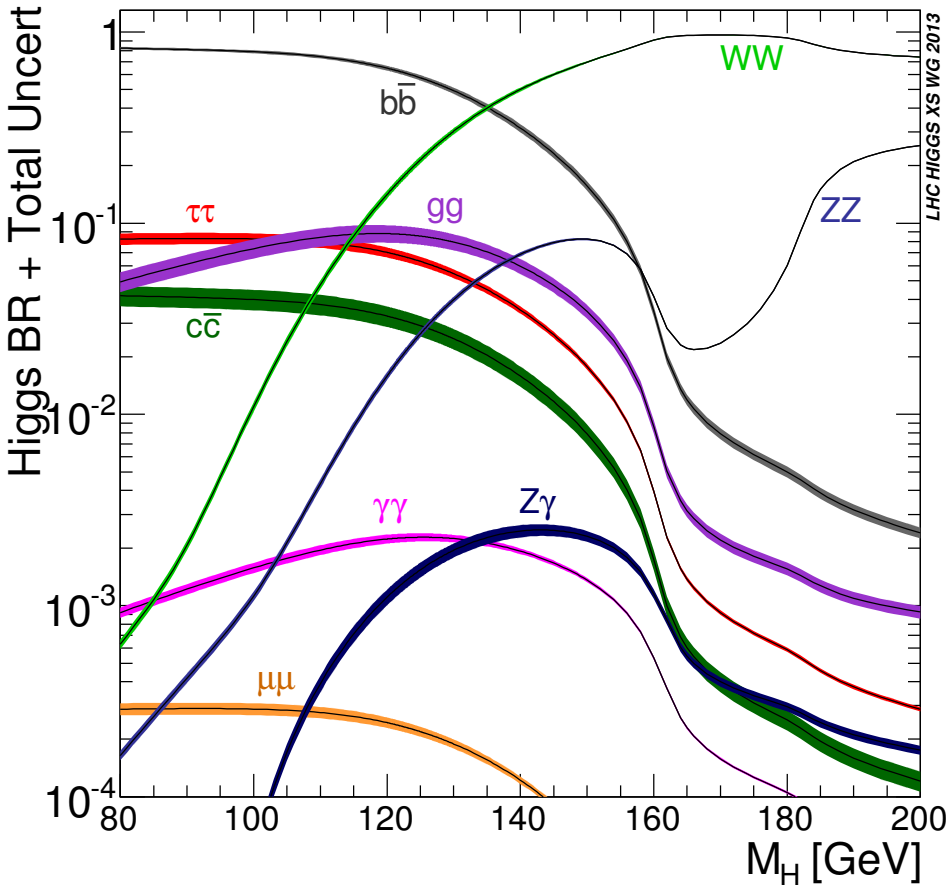


Figure 2.2: Branching ratios for Higgs masses in the range 80 GeV $< m_H <$ 200 GeV. Taken from Ref. [34].

analysis purposes). For larger Higgs masses, the decay $H \to W^+W^-$ is dominant; $H \to ZZ$ and $H \to t\bar{t}$ are prominent for very high m_H.

2.4 The $t\bar{t}H, H \to b\bar{b}$ Channel

Because ATLAS is a general-purpose detector, it is possible to search for a multitude of possible Higgs signals using ATLAS data. These include Higgs production with associated W or Z boson (collectively referred to as "VH") and vector boson fusion ($W^+W^- \to H$ or $ZZ \to H$), which were extensively studied at the Tevatron, as well as gluon-gluon fusion (which has a very high cross section at the LHC because of the parton densities that can be reached at such an energetic hadron collider). Figure 2.3 shows the cross sections for Higgs production via different channels at a proton-

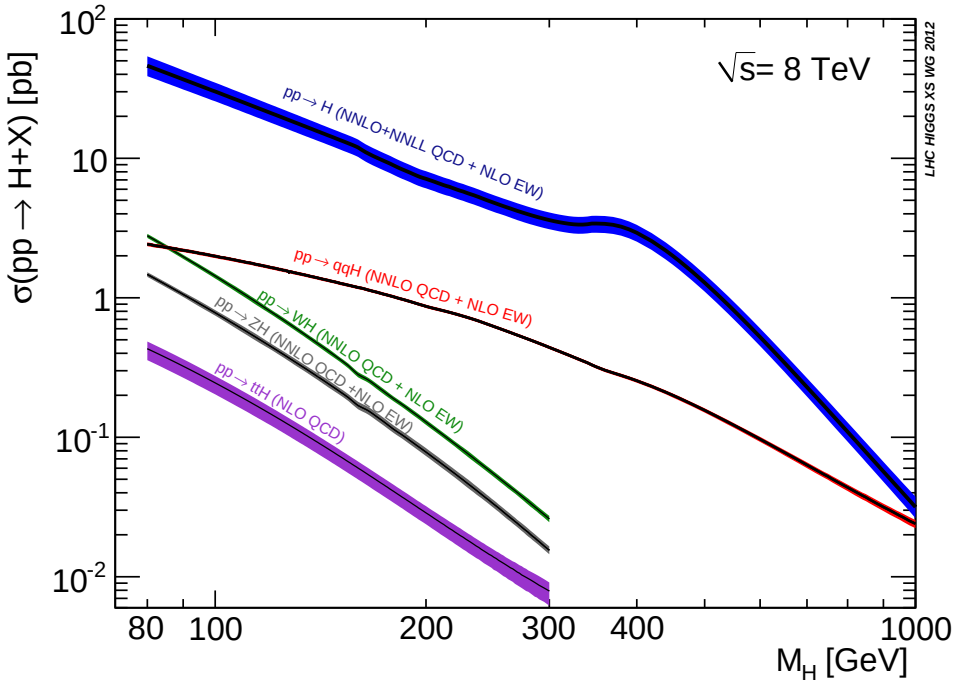


Figure 2.3: Cross sections for different Higgs production modes in a pp collider at $\sqrt{s} = 8\,\text{TeV}$. Taken from Ref. [35].

proton collider for centre-of-momentum (CM) energy[iv] $\sqrt{s} = 8\,\text{TeV}$. Notably, it is not feasible to search for Higgs bosons being produced via gluon fusion in the $H \to b\bar{b}$ channel due to the overwhelming dominance of the QCD multijet backgrounds – potential discovery channels generally involve an associated W or Z boson, or top quark.

One particular gluon fusion channel channel is $gg \to t\bar{t}H$ (with additional contributions from $q\bar{q} \to t\bar{t}H$) [36], *i.e.* gluon-gluon fusion[v] to produce a top-antitop pair, followed by Higgsstrahlung from the top quark (as shown in Figure 2.4), which is the focus of the analysis presented in Chapter 7. In this channel, the top pair will decay via the weak interaction (owing to its extremely large mass, the top quark decays too rapidly to undergo hadronisation): $t \to W^+\bar{b}$ and $\bar{t} \to W^-b$ (decays to other d-type quarks are possible, but since the corresponding CKM matrix elements are very small, these are heavily suppressed).

It is customary to characterise analyses of this channel by the decays of the W bosons originating from the decay of the $t\bar{t}$ system; there are thus three distinct

[iv] CM energy is commonly described in terms of $\sqrt{s}$, where the Mandelstam variable $s \equiv p_1^2 + p_2^2$ for scattering with two initial-state particles having momenta p_1 and p_2.

[v] Or quark-antiquark fusion, but at a pp machine such as the LHC, gluon-gluon fusion is the dominant contribution.

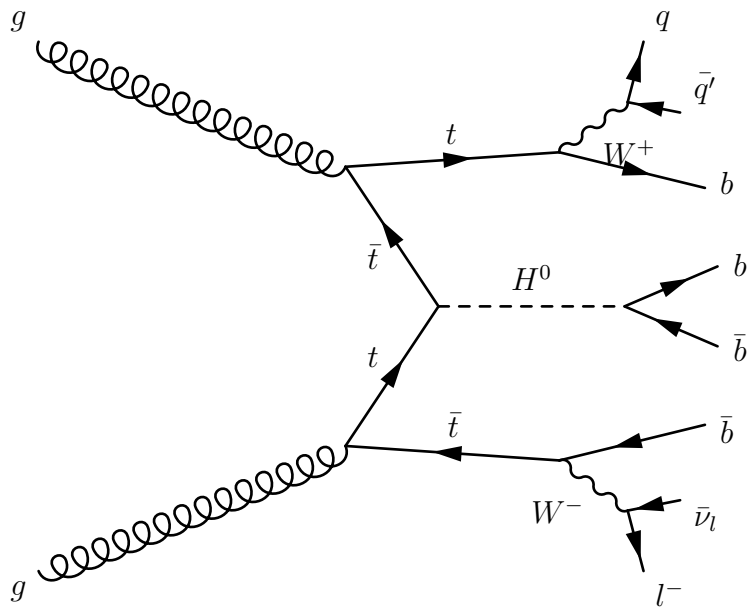


Figure 2.4: Feynman diagram showing one leading-order contribution to the semileptonic (lepton-plus-jets) final state of the $t\bar{t}H, H \to b\bar{b}$ channel

possible kinds of final state – all-hadronic (in which both W bosons decay to quarks), semileptonic (or lepton-plus-jets, in which one W decays to quarks and one decays to a lepton and a neutrino) and dileptonic, in which both W bosons decay to leptons. Of these three, the lepton-plus-jets channel is likely to possess the best discovery potential, since it contains a single hard, isolated lepton, which helps to reduce the background dominance, and it lacks the additional difficulty of reconstructing the $t\bar{t}$ system in the case of two leptonic top decays. Moreover, dilepton events represent only $\frac{1}{9}$ of all $t\bar{t}$ events, since the branching ratio of $W^{\pm}$ bosons to leptons is $\frac{1}{3}$ [8].

2.4.1 The $t\bar{t}H$ inclusive cross section

Theoretical $t\bar{t}H$ cross sections were computed to leading order (LO) as long ago as 1979 [36], however these initial calculations suffered from large theoretical uncertainties because of the dependence of the leading-order cross section on the choice of renormalisation scale μ_{R} for the strong coupling α_s and the choice of proton PDF factorisation scales μ_{F} (*i.e.* the scale choices used to separate hard from soft QCD processes in the parton distribution functions) [37]. Later developments in phenomenology enabled the calculation of these cross sections to next-to-leading order (NLO) (*e.g.* [38]) according to which it was found that the typical ratios of NLO to LO cross sections (the so-called K-factors) were typically in the range 1.14 to 1.22 for the CTEQ6.6 PDF set and 1.05 to 0.98 for MSTW2008 (as m_H increases from

90 to 300 GeV)[vi].

It is notable that, by varying the renormalisation scale μ_{R} and the factorisation scale μ_{F} by a factor of two around their central value μ_0,

$$\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \mu_0 = (2m_t + m_H)/2, \tag{2.19}$$

(where m_t denotes the mass of the top quark and m_H denotes the Higgs mass) the scale dependence was found to be much lower at NLO than at LO – $\mathcal{O}(10\,\%)$ as opposed to $\mathcal{O}(50\,\%)$ – indicating that the NLO predictions were more theoretically robust [37]. The effect of this scale variation on the calculated cross section value was taken to be the contribution to the theoretical cross section uncertainty due to the scale. The other significant contributions to the theoretical uncertainty were the PDFs and α_s.

2.4.2 Discovery prospects

Owing to the very small cross section and the large backgrounds, this channel is not especially promising for Higgs discovery when compared to other modes such as vector boson fusion or associated production with a vector boson (VH). It does, however, provide a unique opportunity to study the Yukawa coupling y_t between the Higgs and the top quark directly; this coupling can otherwise be accessed only indirectly (*e.g.* via the fermion loop in the higher-order process $H \to \gamma\gamma$) or via processes with much lower cross sections, such as $b\bar{b}H$.

2.4.3 Higgs masses and branching ratios in relation to associated production

For Standard Model Higgs bosons within the mass range $115-130\,\mathrm{GeV}$ (*i.e.* the "light Higgs" scenario) the dominant decay mode (as can be seen in Figure 2.2) is $H \to b\bar{b}$. This means that the Higgs will decay to a pair of b-jets (back-to-back within the rest frame of the Higgs) which must be experimentally reconstructed in order to infer the existence of the Higgs (and to calculate its invariant mass, m_H). It therefore follows that the suitability of the $H \to b\bar{b}$ channel for Higgs searches depends largely on the efficiency of b-tagging, *i.e.* the identification of jets originating from b or $\bar{b}$ quarks as such, distinguishing them from "light" jets originating from lighter quarks,

[vi] PDF sets are described in more detail in Chapter 5.

gluons or tau mesons. To reconstruct the Higgs it is necessary to tag three or four of the reconstructed jets as b-jets, in addition to requiring an isolated lepton (from one of the W decays) for triggering purposes and missing transverse momentum (the signature of a neutrino). The Higgs signal would then appear as an excess in the $b\bar{b}$ invariant mass distribution [39].

2.4.4 Backgrounds

In order to discover the Higgs in a given channel it is necessary to discriminate the Higgs signal (*i.e.* the signal-plus-background hypothesis) from the sum of all relevant backgrounds (the background-only hypothesis) with a certain statistical significance (typically 5σ is the threshold confidence level for a discovery). This can be complicated by a number of factors, including the relative cross sections of the signal and backgrounds (*e.g.* if the signal cross section is much smaller than those of the background processes, then there will be accordingly fewer signal events, making it harder to reach a statistically significant result) and the degree to which the backgrounds can resemble the signal (in terms of some variable set which parametrises the events) – if the background processes highly resemble the signal then it is more difficult to distinguish between the signal and random fluctuations of the background.

For the $t\bar{t}H, H \to b\bar{b}$ channel, there are two main kinds of background: the irreducible background, which consists of $t\bar{t}Z, Z \to b\bar{b}$, as well as $t\bar{t}b\bar{b}$ production, and the reducible backgrounds, which contain "light" jets mistagged as b-jets, *e.g.* $t\bar{t}jj$ and $Wjjj$, where j represents a jet that originates from something other than a b quark or its antiparticle. These backgrounds are termed reducible because they act as background processes only by virtue of imperfect b-tagging, and their size can therefore be reduced by implementing improved b-tagging [39, 40]. The rate at which b-jets are tagged as such is referred to as the b-tagging efficiency, and the rate at which light jets are misidentified as b-jets is referred to as the mistag rate. By considering the dependence of the mistag rate on jet variables such as p_{T} and $|\eta|$ it is possible to construct a mistag matrix, which allows one to calculate the probability that a given light jet will be mistagged, as a function of the associated jet variables. This is important in creating simulated data sets (or Monte Carlo), as it is essential to replicate the b-tagging performance of the experiment accurately.

In all experimental regions (as characterised by jet multiplicity and number of b-tagged jets) in the single-lepton $t\bar{t}H, H \to b\bar{b}$ channel, the dominant background

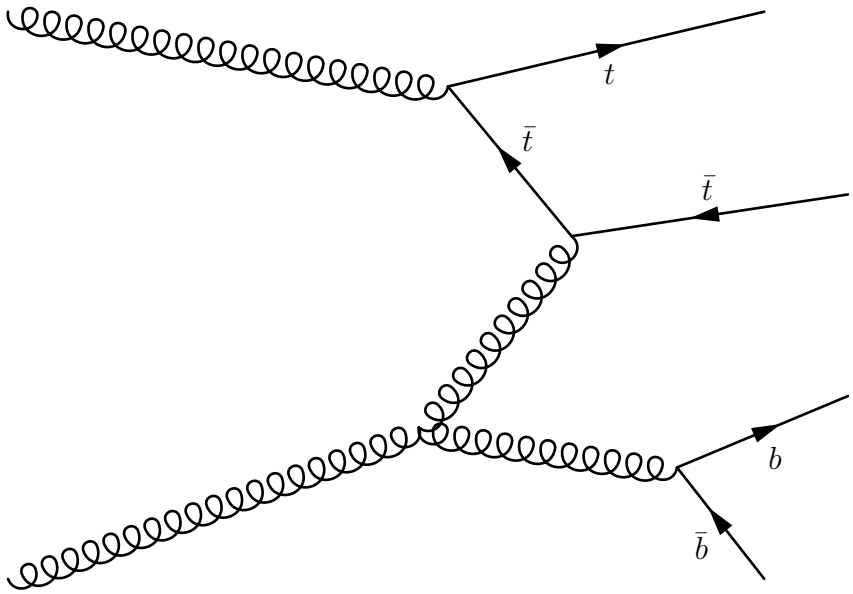


Figure 2.5: Feynman diagram showing a contribution to continuum (background) $t\bar{t}b\bar{b}$ production

contribution is from $t\bar{t}b\bar{b}$ production, *i.e.* continuum production of a b-pair in association with a top pair, a sample Feynman diagram for which is shown in Figure 2.5. One could therefore state the chief objective of a Higgs search in this channel as discriminating *resonant* $b\bar{b}$ production (*i.e.* at the Higgs mass) from *continuum* production (in which case one would not expect to observe a peak at a certain mass in the m_{bb} distribution, for example).

2.5 BSM Higgs Bosons

Many models of physics beyond the Standard Model (BSM) predict the existence of additional Higgs bosons – for example, the MSSM contains five such bosons owing to the additional electroweak degrees of freedom introduced by supersymmetry. Some of these extensions to the Standard Model could in principle give rise to additional contributions to Higgs production in association with the top quark. For example, it is possible that a heavy charged Higgs $H^{\pm}$ with the appropriate couplings to fermions could give rise to final states such as $tt\bar{b}$ and $tt\bar{b}\bar{b}$, as shown in Figures 2.6 and 2.7 respectively.[(vii)] For these heavy-charged Higgs bosons (with $m_{H^{\pm}} > m_t$), associated production of the form $pp \to tbH^{\pm} + X$ dominates [41]. Moreover, for large $m_{H^{\pm}}$ the dominant decay mode is $H^{\pm} \to tb$, although the precise value of the branching

[(vii)] When the notation H^+ is used in the context of charged-Higgs searches in this thesis, the charge conjugate H^- is implicit.

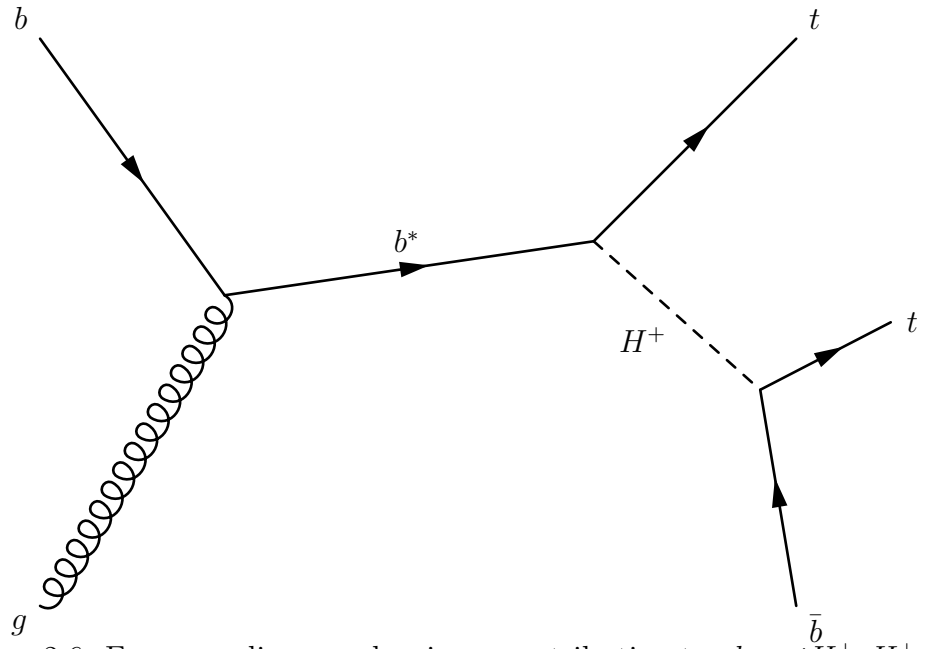


Figure 2.6: Feynman diagram showing a contribution to $gb \to tH^+, H^+ \to t\bar{b}$ in the five-flavour scheme

ratio $\mathcal{B}(H^\pm \to tb)$ depends upon $\tan\beta$; for $\tan\beta = 1$ the branching ratio is of order 1, while for $\tan\beta = 35$ it is around 0.7 to 0.5, thus inducing a model-dependence in the cross-section–branching-ratio product [42]. These $H^\pm$ production cross sections can be computed according to two formalisms, depending on the number of initial-state quark flavours which are considered: the four-flavour scheme neglects b-quarks in the initial state, while the five-flavour scheme permits them [37].

2.6 Higgs Searches at Previous Experiments

The most important previous colliders in relation to Higgs searches were LEP (1989–2000) and the Tevatron (1987–2011, first collisions 1985). The nature of the Higgs searches that were carried out at these colliders are summarised here, together with their results.

2.6.1 Higgs Searches at LEP

LEP, the Large Electron-Positron collider, was an electron-positron synchotron that ran at CERN between 1989 and 2000, preceding the LHC (indeed, the tunnel which now accommodates the LHC was originally built for LEP) [43]. Initially it ran at a centre-of-momentum energy of 91 GeV (*i.e.* the rest mass of the Z boson) to enable studies of the weak neutral current, then it underwent an upgrade in 1995 to

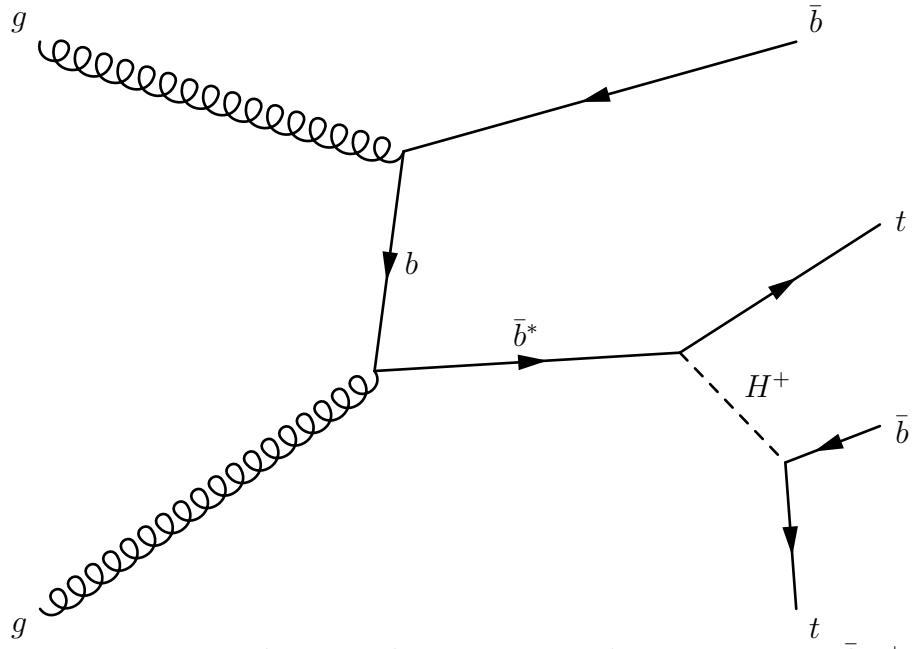


Figure 2.7: Feynman diagram showing a contribution to $gg \to t\bar{b}H^+, H^+ \to t\bar{b}$ in the four-flavour scheme

enable running above the threshold of W pair production (up to a maximum CM energy of 209 GeV), primarily to allow for precision studies of the charged-current sector of the weak interaction [44]. LEP is particularly known for its extensive work in making precision measurements of the W and Z boson masses [45, 46], in addition to confirming the Standard Model prediction of three neutrino flavours. The experiment comprised four detectors (ALEPH, DELPHI, OPAL and L3), each of which was involved in extensive Higgs searches, including more exotic models such as two Higgs doublet models, as well as fermiophobic Higgs and invisible Higgs decay modes. The final limit on m_H that was set by the searches performed at LEP is shown in Figure 2.8.

Higgs searches at LEP focussed primarily on the Higgsstrahlung process $e^+e^- \to ZH$, with a kinematic threshold of $\sqrt{s} = m_H + m_Z$. There was some additional sensitivity gained via the vector boson fusion channels $e^+e^- \to e^+e^-H$ (*i.e.* ZZ fusion) and $e^+e^- \to \nu_e\bar{\nu}_eH$ (*i.e.* WW fusion). At the Higgs masses that LEP was sensitive to, the dominant decay modes would have been $H \to b\bar{b}$ and $H \to \tau^+\tau^-$. The collider's CM energy was increased every year during the second run in order to acquire sensitivity to higher Higgs masses. In addition, Higgs sensitivity was increased by maximising the integrated luminosity through statistically combining data from the four detectors. In total, the LEP Higgs working group had 2461 pb^{-1} of data from the LEP2 run, all of which was collected at CM energies of $189 \leq \sqrt{s}/\text{GeV} \leq 209$ [47].

The combined LEP analysis yielded a lower bound of 114.4 GeV on the Higgs mass (at the 95 % confidence level), with a median expected limit of 115.3 GeV [48]. In addition, precise measurements of electroweak observables (which depend logarithmically on the Higgs mass via loop corrections) allowed LEP to set an upper bound of $m_H < 186\,\text{GeV}$ at the 95 % confidence level. The LEP electroweak working group gave a final result of $m_H = 91^{+58}_{-37}\,\text{GeV}$ from electroweak observable fits to data [47]. LEP Higgs searches also found a small excess (of 1.7 standard deviations significance compared to the background-only hypothesis)[(viii)] for a Higgs candidate with mass 115 GeV, lending some credence to the possibility that the Higgs mass lay just above the region within which the LEP analysis had sufficient sensitivity to carry out an exclusion.

2.6.2 Higgs searches at the Tevatron

The Tevatron was a proton-antiproton collider at the Fermi National Accelerator Laboratory (Fermilab) which had two detectors, CDF and DØ, both of which were involved in Higgs searches. As a hadron collider, the Tevatron was capable of colliding at much higher centre-of-momentum energies (up to 1.96 TeV) than lepton colliders such as LEP. Nonetheless, the interaction dynamics of hadron colliders are far more complex due to the composite nature of protons and antiprotons. These factors mean that more physics can be accessed at hadron colliders (making them highly suitable for discoveries), yet precision measurements (and analyses in general) are much more difficult. The Tevatron is known primarily for its discovery of the top quark in 1995, and it was hoped that during its second run, after an upgrade had increased its instantaneous luminosity by an order of magnitude, it might be able to discover the Higgs, or at least exclude it within certain regions of phase space (depending on m_H).

At the Tevatron, the Higgs production process with the highest cross section was gluon-gluon fusion mediated by top and bottom quark loops ($\sigma \approx 0.7\,\text{pb}$ for $m_H = 120\,\text{GeV}$) [49], although the backgrounds are very large. Because of this, the most interesting discovery channels were Higgs production with an associated W or Z boson, where the latter decays leptonically (because decays to hadronic jets result in less clean final states). Such channels are $WH \to l\nu b\bar{b}$, $ZH \to \nu\bar{\nu}b\bar{b}$ and $ZH \to l^+l^-b\bar{b}$. Combining these channels with $H \to W^+W^-$ gives an expectation of approximately 500 Higgs bosons produced for any mass point m_H satisfying

[(viii)] A more detailed discussion of statistical significance is given in Chapter 6.

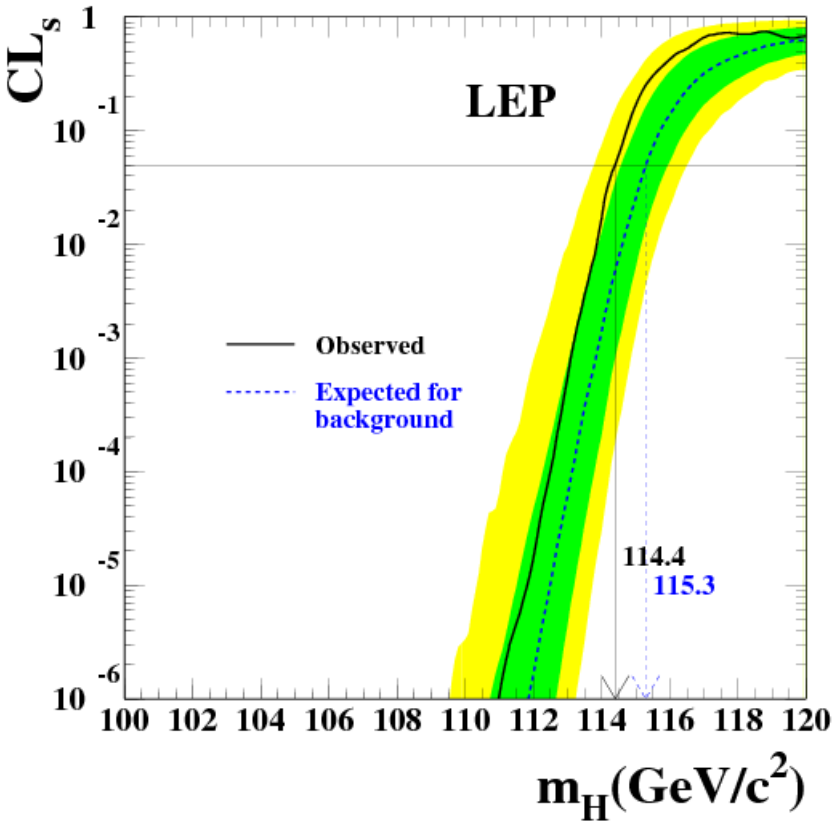


Figure 2.8: Results of the LEP Higgs searches, showing a lower bound $114.4\,\text{GeV} < m_H$. The statistical concepts underlying the production of such exclusion plots are explored in detail in Chapter 6. Figure taken from Ref. [47].

$114\,\text{GeV} < m_H < 185\,\text{GeV}$ [49]. The Tevatron was ultimately able to exclude the Higgs in the mass regions 100–108 GeV and 156–177 GeV to the 95 % confidence level, as illustrated in Figure 2.9.

2.6.3 Higgs searches at the LHC

The LHC is the successor machine to the Tevatron at the high-energy frontier of experimental particle physics; from the very beginning of the program one of the most important design objectives was to maximise the machine's ability to study the Higgs. To ensure a maximal total Higgs cross section, the machine would need to run at a much higher CM energy than any of its predecessors, and it would need to deliver a far greater luminosity. Specifically, the LHC's Higgs program centres on its two general-purpose detectors, ATLAS and CMS – these are the largest detectors yet constructed for a particle physics experiment.

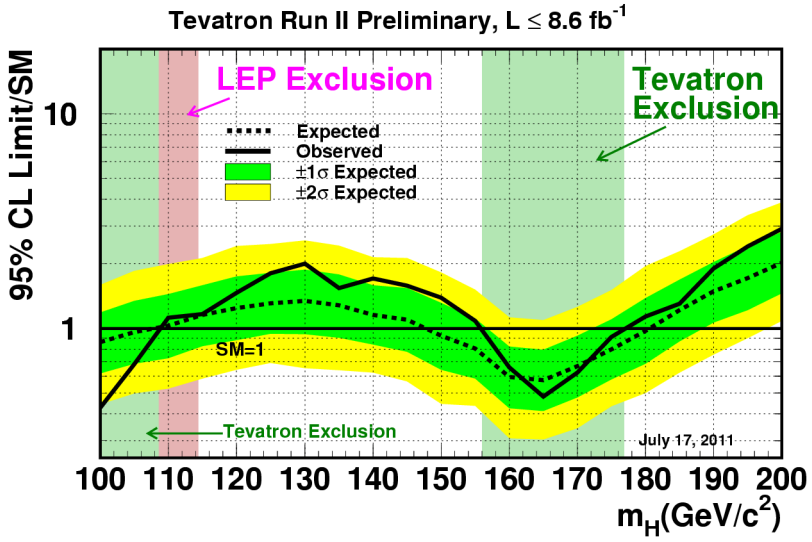


Figure 2.9: Results of the Tevatron Higgs searches, showing the excluded mass regions 100 GeV $< m_H <$ 108 GeV and 156 GeV $< m_H <$ 177 GeV. The LEP exclusion (as in Figure 2.8) is shown for comparison. Figure taken from Ref. [30].

The LHC's first run (commonly termed "Run 1") began in 2009, with extensive data-taking beginning in 2011; during this year ATLAS accumulated 4.8 fb^{-1} of proton-proton collision data at a centre-of-momentum energy of 7 TeV. The most important Higgs analyses using this data were those involving Higgs decays to bosons, specifically $H \to ZZ^* \to llll$, $H \to \gamma\gamma$ and $H \to WW \to e\nu\mu\nu$ [50–53]. In July 2012, the ATLAS and CMS collaborations jointly announced their discovery of a new scalar boson consistent with the Standard Model Higgs, as observed in these channels, using the 7 TeV dataset in addition to collision data from the beginning of the 8 TeV run that occurred during 2012 [54]. It is notable that Higgs decays to fermions have not yet been observed – ongoing analyses focus specifically on $H \to \tau\tau$ and $H \to b\bar{b}$, as these have the highest branching ratios for a light Higgs; the analysis presented in this thesis is a search for $H \to b\bar{b}$.

2.7 Summary

The theoretical underpinnings of the Standard Model have been introduced and explored, with particular regard to electroweak physics. The theoretical motivation for spontaneous electroweak symmetry breaking has been explained in relation to

Higgs physics and Higgs searches at previous and current experiments, including at the LHC. The relationship between the Higgs boson and the top quark has been discussed, together with the status and prospects of ongoing Higgs searches in the channel $t\bar{t}H, H \to b\bar{b}$. The theoretical extension of the Standard Model by the incorporation of supersymmetry to give the MSSM has been explored, together with the nature of the Higgs sector in the MSSM and in particular the possibility of charged-Higgs discovery in the channel $gg \to tH^{\pm}, H^{\pm} \to tb$.

Chapter 3

The Large Hadron Collider

> "Nil sine magno vita labore dedit mortalibus."
>
> — Horace, *Sermones*

> "... omnia praeclara tam difficilia quam rara sunt."
>
> — Spinoza, *Ethica*

The Large Hadron Collider (LHC) [2, 3], situated at the European Organization for Nuclear Research (CERN) in Geneva, is the world's largest particle accelerator to date, occupying a 27 km tunnel which formerly housed the LEP accelerator, traversing the French-Swiss border at an typical depth of around 100 m with an inclination of approximately 1.4 % with respect to the horizontal. It has been designed to produce extremely high-energy collisions, chiefly proton-proton but also with heavy ions (Pb–Pb and p–Pb); it supplies collisions to four large experiments which collectively support an extremely large and diverse physics program covering most areas of ongoing research in contemporary high-energy physics. The LHC constitutes the largest component of the CERN accelerator complex, as depicted in Figure 3.1. The nominal (*i.e.* design specification) values of some important LHC parameters are listed in Table 3.1.

Notably, the machine has been designed to deliver an extremely high instantaneous luminosity, $L = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ – this is a measure of the rate at which collision data is produced, in terms of events per square centimetre of cross section σ per second. The event rate is then simply the product $N = L\sigma$, and the integrated luminosity is just $\mathcal{L} = \int L\,\mathrm{d}t$. Most typically $\mathcal{L}$ is measured in terms of events per *barn* (symbol: b) of cross section, where $1\,\mathrm{b} \equiv 10^{-24}\,\mathrm{cm}^2$. Thus one might speak of a detector as having accumulated a certain number of inverse femtobarns of data – *e.g.* $5\,\mathrm{fb}^{-1}$,

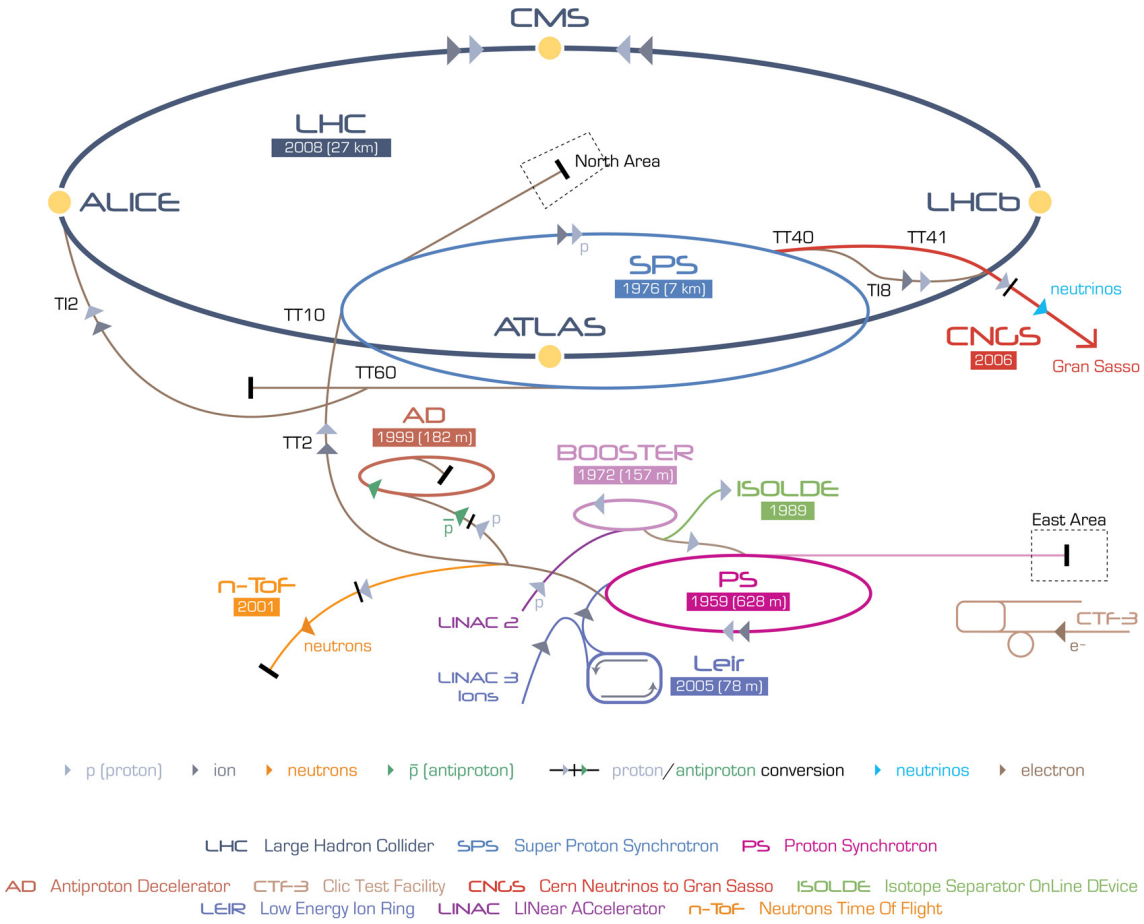


Figure 3.1: Diagram of the full CERN accelerator complex, including the LHC. Taken from Ref. [55].

Parameter	Nominal Value
Energy per proton (maximum)	7 TeV
Energy per beam (maximum)	366 MJ
Luminosity	$10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
Beam lifetime (maximum)	28 h
Maximum number of bunches	2808
Protons per bunch	1.15×10^{11}
Collisions per bunch crossing (mean)	19
Bunch separation (minimum)	25 ns
Integrated cross section (per machine cycle)	100 mb
Proton energy loss per turn	6.7 keV
Temperature	1.9 K
Beam pipe vaccuum	$10^{-8} - 10^{-9}$ Pa

Table 3.1: Table listing LHC parameters with their nominal values (for *pp* collisions) [2, 3].

equivalent to stating that one expects to observe 5 events per femtobarn of collision cross section in the data.

For a hadron collider, as compared to a lepton collider, luminosity and beam-dimension measurement are nontrivial (in the case of lepton colliders such as LEP, the Bhabha scattering process $e^+e^- \to e^+e^-$ provides an adequate means of monitoring luminosity [56]). Luminosity monitoring techniques for hadron colliders include beam-gas imaging [57] and Van Der Meer scans [58, 59]. The luminosity can also be computed according to the relation

$$L = \frac{N_1 N_2 n_1 n_2 f \gamma}{4\pi\sigma_x\sigma_y} F(\theta_c), \tag{3.1}$$

where N_1 and N_2 are the numbers of particles in each bunch[(i)], n_1 and n_2 are the numbers of bunches in each beam, f is the frequency of revolution, γ is the Lorentz factor and σ_x and σ_y are the effective dimensions of the beam as measured using Van Der Meer scans [60]. $4\pi\sigma_x\sigma_y$ is the effective area of the collision. Finally, $F(\theta_c)$ is a factor representing the effect of the crossing angle – to avoid so-called parasitic

[(i)] When collisions are occurring, the protons are not distributed uniformly throughout the ring, but are concentrated in packets called *bunches*. These bunches are further organised into so-called *bunch trains*.

collisions (*i.e.* collisions outside the desired interaction region), the beams do not collide head-on [58].

The "size" of a beam may be quantified in terms of the transverse emittance ϵ and the amplitude function (or "betatron function") at the interaction point β^*, defined as

$$\beta^* = \frac{\pi\sigma^2}{\epsilon}, \tag{3.2}$$

where σ is the cross-sectional width of the bunch; β^* has dimension of length. The amplitude function reflects the "squeezing" effect of the quadrupole magnets on the beam's cross-sectional shape; typically $\mathcal{L} \propto \frac{1}{\beta^*}$ [61]. Roughly speaking, the purpose of the quadrupole magnets at the interaction points is to compress the bunches so as to minimise the amplitude function in these locations and hence ensure the greatest possible density of collisions.

The transverse emittance has dimensions of length, and it reflects how tightly "bunched together" the protons are – a function of how the beam is prepared at the point of injection into the accelerator complex. In order to maximise the luminosity, it is desirable to minimise the emittance of the beam, as a low-emittance beam is one in which the protons are confined to a small region and have similar momenta; they will not disperse over time.

A further important design consideration is *pileup* [62], which can take the form of in-time pileup or out-of-time pileup. In-time pileup denotes the number of interactions per bunch crossing, which varies with the size of the effective interaction area (*i.e.* tighness of beam focus) and the number of protons per bunch. Out-of-time pileup, by contrast, describes additional interactions that occur due to protons originating from previous bunch crossings. It is important to contrast the short bunch spacing (25 ns nominal) with the readout times for some of the ATLAS detector subsystems – in the case of the electromagnetic calorimeter, up to 600 ns. Because of this difference, decreasing bunch spacing would also increase out-of-time pileup: the detector would be less able to match any given particle track to a specific bunch crossing.

3.1 Machine Superstructure

The LHC ring is subdivided into eight independent sectors, each of which runs from the centre of one insertion (a straight region of the accelerator) to the centre of the

next insertion. These insertions serve as regions along the ring where beams can be injected, dumped or collided, and they allow access to the machine from the surface – the four experiments are located at such points (ATLAS at Point 1 beside the main CERN campus and CMS diametrically opposite at Point 5, with ALICE and LHCb adjacent to ATLAS at Points 2 and 8 respectively). Between the insertions, the accelerator curves along arc regions, where dipole magnets bend the beams.

The dipole magnets used in the accelerator have a bend radius of approximately 2.8 km, and they produce a magnetic field of strength 8.3 T. In order to generate such a strong magnetic field, extremely large electric currents (around 12 kA) are needed, necessitating the use of superconducting electromagnets – these are cooled to approximately 1.9 K to maximize their performance. The main ring of the LHC contains 1232 of these 15-metre cryodipole magnet segments.

The high rate at which the beam deposits its energy into the magnet system in the form of synchotron radiation – 6.7 keV per proton per turn, a heating power equivalent to 3.6 kW across the ring – necessitates the use of specialised cryogenic technology. Liquid helium is used for this purpose owing to the excellent heat transfer properties it possesses in its superfluid state, allowing several kilometres of dipole magnets to be cooled down to a few kelvins of absolute temperature with very little temperature fluctuation. The LHC's superconducting magnets are immersed in this supercooled helium at atmospheric pressure, and the low temperature is maintained by replenishment of the helium. As liquid helium is a comparatively scarce and expensive resource, the initial cooling cycle of the LHC is carried out by vaporising approximately 12 000 m^3 of liquid nitrogen – only once the machine has been cooled to the boiling point of nitrogen does the following stage of cooling by immersion in the superfluid helium bath commence. The LHC maintains an inventory of about 700 m^3 of liquid helium for this purpose [63].

Over each interaction region, for a distance of about 120 m the two beams are contained in a single pipe, with their separation minimised at the interaction point in order to maximise collisions there. Asymmetric triplet magnets are used to "squeeze" the two beams so that they both focus on an extremely small region – the interaction point, where the beams intersect with a typical crossing angle θ_c of approximately 300 µrad.

The injection cycle comprises several distinct phases. The initial hydrogen atoms are ionised to produce protons, which are then accelerated by Linac2 to an energy of 50 MeV and injected into the Proton Synchotron Booster. Then they are boosted to 1.4 GeV and injected into the Proton Synchotron (PS) itself, which raises the

energy further to 25 GeV. The Super Proton Synchotron (SPS) then prepares the beam for injection into the LHC, raising the beam energy to 450 GeV and injecting the protons into the LHC via two separate beampipes, one running in each direction. The LHC then carries out the final phase of the acceleration, ramping its magnets once the beam has been injected and raising the beam energy to the TeV scale (up to a nominal maximum of 7 TeV per beam) and enabling collisions at the interaction points.

3.2 The Main LHC Experiments

Once the two beams have reached the desired energy inside the LHC, they may be steered onto collision trajectories by dipole magnets located close to the four interaction points. These magnets squeeze the beams as described above, minimising their cross-sectional widths and introducing a crossing angle between them. This process occurs inside each of the major experiments, the goals of which are summarised below.

3.2.1 ALICE

Standing for "A Large Ion Collider Experiment", ALICE has been designed with a view to studying heavy-ion (*e.g.* lead-lead or proton-lead) collisions, including phenomena such as quark-gluon plasma – a state of extremely energetically dense matter in which the asymptotic freedom of partons at high-Q^2 is realised. This is believed to replicate the conditions that prevailed in the very early universe, before the mean temperature reduced to a level at which hadrons could form [64]. ALICE became the first LHC experiment to publish an analysis of proton-proton collision data in November 2009 [65].

3.2.2 ATLAS

ATLAS is a large general-purpose detector which has been designed to study an extremely wide range of phenomena within the standard model and beyond. Its most important physics objectives include Higgs searches, studies of the properties of the top quark, and direct searches for SUSY and other BSM phenomena. It is an hermetic detector with nearly full 4π sr solid-angle coverage of its interaction point, as required in order to be able to fulfil its physics objectives [4].

The ATLAS detector is described in detail in Chapter 4.

3.2.3 CMS

The Compact Muon Solenoid, or CMS, is the LHC's other large general-purpose detector. It supports a physics program as varied as that of ATLAS, including Higgs physics, searches for BSM phenomena such as supersymmetric particles and extra dimensions, and exotic physics.

The ATLAS and CMS collaborations were intended from the outset to conduct very similar research programs in fundamentally independent ways, with separate researchers, independent analyses and different detector technologies. Each analysis is performed separately at each experiment until a final, public result is obtained, at which point the results can be directly compared and statistically combined with one another. Moreover, this approach provides a unique and valuable check on results from either of the two experiments [5, 66, 67].

3.2.4 LHCb

LHCb ("b" representing the bottom, or beauty quark) is the smallest of the four main LHC experiments. It is an asymmetric detector concentrated in the forward region, which specialises in the study of heavy-flavour physics, such as $\mathcal{CP}$-violation (*i.e.* the asymmetry between matter and antimatter) and measurements of rare decays of hadrons containing b and c quarks [68].

Chapter 4

The ATLAS Detector

> "Und laß dir raten, habe
> Die Sonne nicht zu lieb und nicht die Sterne;
> Komm, folge mir ins dunkle Reich hinab!"
>
> — Goethe, *Iphigenie auf Tauris*

4.1 ATLAS: Background and Overview

4.1.1 Introduction

The ATLAS detector[(i)] is, along with CMS, one of the LHC's two general-purpose detectors, meaning that it is intended to be capable of studying as wide a range of phenomena as possible, potentially opening up the entirety of TeV-scale physics to exploration; this approach contrasts with that of more specialised detectors such as LHCb, which has a very specific focus on heavy-flavour physics. The basic cylindrical structure of the ATLAS detector is illustrated in Figure 4.1. It is intended to be a hermetic detector, meaning that it aims for complete 4π sr-solid-angle coverage of the interaction point; it is symmetric around the beamline and on either side of the interaction point. Enclosing a volume of approximately $20\,000\,\mathrm{m}^3$, ATLAS is by far the largest high-energy physics experiment by volume constructed to date, although CMS is more massive [4].

ATLAS is located at Point 1 on the LHC ring, adjacent to the main CERN site at Meyrin and approximately 100 m underground. It measures 46 m long, with a diameter of 25 m and a total mass of 7000 t; it contains approximately 3000 km

[(i)] "ATLAS" is a contrived and nested acronym, standing for **A** **T**oroidal **L**HC **A**pparatu**S**.

of cable, and represents the culmination of 20 years of planning, development and construction. The experiment was first proposed in the late 1980s, and construction occurred between 1997 and 2008. The LHC ran its first beam in September 2008, but technical problems with the accelerator magnet system necessitated repairs, which delayed the first proton-proton collisions until November 2009 [65]. These early collisions, occurring at a centre-of-momentum energy of about 1.8 TeV, broke the energy record that had been held by Fermilab's Tevatron for the previous 22 years.

The design requirements of ATLAS are exacting owing firstly to the capabilities of the LHC accelerator, which has a design maximum centre-of-momentum energy of 14 TeV, a nominal bunch spacing[(ii)] of only 25 ns, with a mean of 19 events per bunch crossing, and an instantaneous luminosity[(iii)] of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. ATLAS was designed to be sensitive to signals for physics processes with much smaller cross-sections and far larger backgrounds than any that had previously been observed. Because of the wide variety of physics studied by ATLAS, the detector's subsystems have been optimised according to a wide range of requirements:

- The high luminosity and energy, and the large flux of particles, mean that radiation hardness requirements on all detector components are stringent.
- The high collision rates necessitate very fast electronics to allow effective processing of events and sufficiently rapid triggering.
- Very fine granularity is needed throughout the detector because of the high multiplicity of particles per event and because of high event overlap (*i.e.* several simultaneous events to disentangle).
- Acceptance must be as extensive in η and ϕ as possible; in principle, full azimuthal coverage over the whole observable range of pseudorapidity is desirable to ensure that no high-energy particles escape detection.
- Trigger rate: trigger on high-p_{T} particles to optimise kimenatic efficiency and background rejection.
- Muons must be identified according to their charge with good accuracy, and their momenta must be measured with high precision.

[(ii)] The bunch spacing is the time delay between consecutive bunches of protons as they circulate around the collider.

[(iii)] The instantaneous luminosity is the rate at which events are accumulated per unit of scattering cross-section area.

- Vertex reconstruction; efficient tracking of high-p_{T} leptons; ID of electrons, positrons and photons, and measurements of their energies; ID of τ leptons and heavy flavours; measurement of jet energies and $E_{\mathrm{T}}^{\mathrm{miss}}$.

4.1.2 Coordinates

Points within the detector are conventionally identified with reference to a standard coordinate system; ATLAS uses a system of spherical polar coordinates of the form (r, θ, ϕ), where the interaction point is taken as the origin of the system. The beamline is taken as the z-axis, perpindicular to the x-axis (running from the interaction point towards the geometric centre of the LHC) and the y-axis (running from the interaction point vertically upwards). The azimuthal angle $\phi = \tan^{-1}(y/x)$ is measured in the x-y plane around the beamline, and the polar angle $\theta = \cos^{-1}(y/x)$ is the angle measured with respect to the positive direction of the beamline (*i.e.* $+z$), although typically it is more useful to employ a Lorentz-invariant measure of angle, the pseudorapidity η, defined by

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{4.1}$$

This defines a measure of angle which is unchanged by Lorentz boosts along the beam line, and hence does not vary between the CM frame, the laboratory frame, and the reference frames of individual particles travelling along the beam line; significantly, the rate of particle production is approximately constant as a function of η. The pseudorapidity is defined such that it is zero for trajectories perpendicular to the beamline, and approaches positive (negative) infinity as a trajectory becomes parallel with the positive (negative) beam direction. Typically, the high-η region (*i.e.* small angular separation from the beamline) is spoken of as the *forward* (or, equivalently for a symmetric machine, the *backward*) region of the detector.

Another common measure of angle is the rapidity y, defined by

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) = \ln\left(\frac{\sqrt{m^2 + p_{\mathrm{T}}^2\cosh^2\eta} + p_{\mathrm{T}}\sinh\eta}{\sqrt{m^2 + p_{\mathrm{T}}^2}}\right). \tag{4.2}$$

Notably, the definition of y used by particle physicists generally uses longitudinal momentum p_z instead of the absolute value of the three-momentum $|\boldsymbol{p}|$.

Finally, it is common to measure the total angular separation of pairs of coordinates

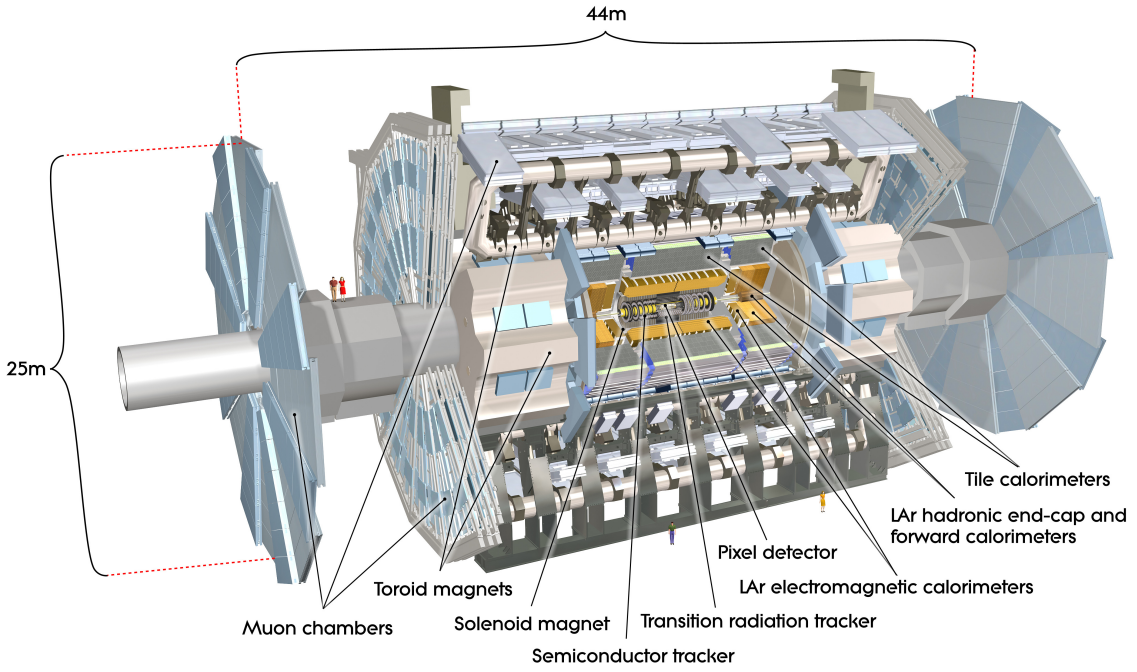


Figure 4.1: Cutaway of the ATLAS detector in profile, with major subsystems indicated. Taken from Ref. [69].

(for instance, it might be desirable to study the angular separation of different pairs of jets, to measure which combination yields the minimal separation) by means of the quantity ΔR, defined by

$$\Delta R = \left| \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \right| . \tag{4.3}$$

4.2 Detector Structure

ATLAS has a complex substructure, as shown in Figure 4.1. It comprises several distinct layers of detectors, each with a specific purpose. From the interaction point, particles radiate outwards through these layers, encountering firstly the inner detector, which is responsible for tracking and particle identification, then the calorimeters, which measure the energies of charged and strongly-interacting particles, and then the muon system, which identifies muons and measures their momenta. Moreover, the successful operation of the experiment as a whole depends on the efficient exchange of information between these different detector technologies and a system which is responsible for making rapid decisions about which events to discard as unlikely to contain interesting physics (this decision process is termed *triggering*).

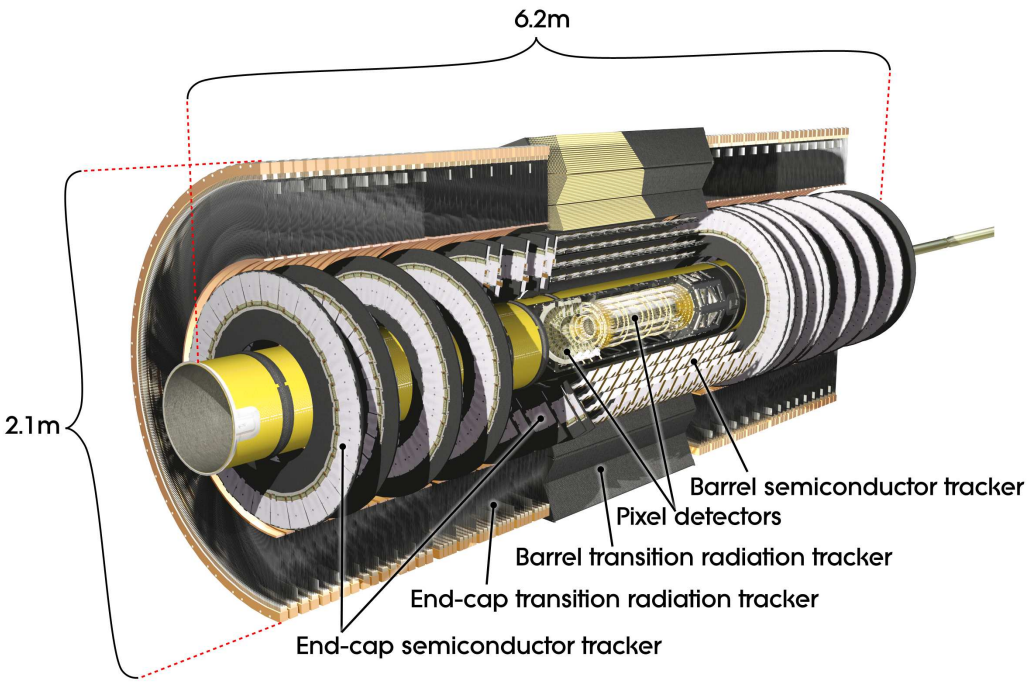


Figure 4.2: Diagram of the ATLAS Inner Detector. Taken from Ref. [70].

This framework is referred to as TDAQ, or Triggering and Data AcQuisition.

These distinct subsystems will each be explained in turn in the following sections.

4.2.1 The Inner Detector

The ATLAS inner detector (ID), as illustrated in Figure 4.2, comprises three subsystems: the silicon pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT). The main function of the inner detector as a whole is to make precision measurements of the momentum of charged particles; it also satisfies several requirements on, for example, electron and photon identification, as well as secondary vertex reconstruction and the identification of jets originating from b quarks. These requirements determine the granularity of the detector subsystems, as well as the range in η and r which the detector must cover. The ID has a length of 6.2 m and is 2.1 m in diameter; its angular coverage is within the region $|\eta| \leq 2.5$. It is immersed in a 2 T magnetic field, generated by a solenoid.

The silicon pixel detector is the closest detector to the beamline – it is also the detector system with the highest granularity, with minimum pixel dimensions of 400 µm × 50 µm, and over 80 million readout channels, designed to make precision measurements over the full acceptance of the detector. As the first detector subsystem that particles from the interaction point encounter, the pixel layers play

an important role in vertex reconstruction for particles which decay over relatively short distances (of order millimetres to centimetres), such as B hadrons – this is the objective of the innermost part of the pixel detector, the vertexing layer, at a typical distance of 5 cm from the beamline. There are two further barrel layers at radii of 9 cm and 12 cm, and three disks on each side of the barrel, which extend from 9 cm to 15 cm in radius; the intention is that each track should cross three layers of pixel modules. Each module features 46 080 pixels (16 arrays of 18×160 pixels, each with its own readout chip); each barrel layer comprises 1456 modules and each endcap layer comprises 288 modules. The radiation hardness requirements are exacting, with each 62.4 mm by 21.4 mm module needing to withstand a total dose of absorbed ionising radiation of at least 300 kGy and an integrated neutron flux of at least $5 \times 10^{14}\,\mathrm{cm}^{-2}$ over its operational lifetime.

Further from the interaction point is the semiconductor tracker, which consists of layers of silicon microstrip detectors. The pixel and strip detectors are designed with the intention that any particle leaving the interaction point should cross at least three pixel layers and four strip layers; multiple such crossings are desirable, since each interaction of a particle with a detector layer gives an additional data point for particle ID and event reconstruction. Each SCT module is subdivided into four sensors, each of which measures 63.6 mm by 64.0 mm and features 768 readout strips. The SCT is divided into three regions; a barrel (the central region) and two endcaps; in the barrel the detector modules are arranged in a cylindrical configuration with four concentric layers, whereas in the endcaps, they are arranged as discs centred on the beam axis, with the strips oriented radially. Throughout the SCT, consecutive layers of strips are oriented relative to one another at a small stereo angle of 40 µrad in order to enable z-coordinate measurement. The SCT contributes to measurement of momentum, vertex position and impact parameter, and it plays an important role in identifying jets; it has some 6.3 million readout channels.

The final component of the inner detector is the transition radiation tracker, which is intended to provide drift time measurements and continuous tracking (typically around 30 hits per track), with a view to improving momentum resolution and pattern recognition. Measurements of transition radiation photons compliment the EM calorimeter in electron and positron identification.

The TRT comprises radial tungsten-rhenium wires of 30 µm diameter, each of which is isolated in its own volume of nonflammable gas inside a "straw" of 4 mm diameter; the TRT includes a total of 3 m^3 of Xe(70 %)-CO_2(27 %)-O_2(3 %) gas mixture. Each

straw provides a spatial resolution of 170 μm, together with two distinct energy thresholds. This allows the detector to discriminate between hits which pass only the lower threshold (tracking hits) and those which pass the higher threshold (transition radiation hits).

Inside the TRT central barrel there are 50 000 straws, contained in modules of between 329 and 793 straws each, which can be up to 1.44 m in length; the straws cover radial distances between 56 cm and 107 cm from the interaction point. These straws are divided along the middle of the barrel into two halves[(iv)], each of which is read out separately to reduce occupancy, giving a total of 100 000 channels which are read out from the barrel. The two endcaps contain 320 000 straws in total, which are radially configured in concentric wheels (18 per endcap). In total, then, the TRT provides data from 420 000 channels.

In general it is possible to express the resolution of a given track parameter X as a function of p_{T} according to the formula

$$\sigma_X\left(p_{\mathrm{T}}\right) = \sigma_X\left(\infty\right)\left(1 \oplus p_X/p_{\mathrm{T}}\right) , \tag{4.4}$$

where $\sigma_X\left(\infty\right)$ is the asymptotic resolution (*i.e.* the limit as p_{T} is taken to infinity) and p_X is the momentum value at which the contributions of the intrinsic and multiple-scattering terms are equal for the parameter X. The symbol $\oplus$ denotes addition in quadrature [71].The total performance of the tracking as a function of p_{T} is given by the formula

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = 0.05\,\% \frac{p_{\mathrm{T}}}{\mathrm{GeV}} \oplus 1\,\% , \tag{4.5}$$

where the first term represents the sampling resolution and the second term represents the effect of multiple Coulomb scattering – this effect ultimately constrains the resolution of the detector.

4.2.2 The Calorimeters

The detector employs two kinds of calorimeter, electromagnetic and hadronic; their structure is illustrated in Figure 4.3. The calorimeters are designed to measure the energies of particles which interact via the electromagnetic and strong interactions

[(iv)] The central 80 cm portion of each straw is inactive over the first six layers to further reduce occupancy.

respectively. In each case, a particle entering the calorimeter triggers a particle shower, *i.e.* the cascade of secondary particles produced as the particle from the hard-scattering process decays or hadronises. Notably, the measurements made by calorimeters are typically destructive, meaning that the energy of the particles is completely absorbed as they by measurement.

In general, the energy resolution of a calorimeter can be parameterised as

$$\frac{\sigma_E}{E} = \sqrt{\frac{a^2}{E} + \frac{b^2}{E^2} + c^2}\,, \tag{4.6}$$

where a is the stochastic term, b is the noise term and c is a constant term which accounts for non-uniformities and miscalibrations [72]. In most instances, the stochastic and constant terms are dominant, and these are given for each subsystem in the following sections.

Electromagnetic showers are typically produced by high-energy photons, electrons and positrons; above a certain threshold energy (typically on the order of megaelectronvolts) the dominant means by which photons interact with matter is via pair production (as opposed to Compton scattering and the photoelectric effect, which are dominant for softer photons). As they enter matter, high-energy photons convert into electron-positron pairs; these electrons and positrons then emit further photons as they interact with the nuclei and electrons of the detector (*i.e.* Bremsstrahlung). This process continues until the particles in the shower have depleted their energy below a critical threshold E_c, at which point the shower stops.

Hadronic showers, in contrast, typically involve the decay and hadronisation of strongly-interacting particles, with the energy of the initial particle typically being divided between its daughter particles and multiparticle interactions (such as soft pion production) in an approximately equal fashion. Hadronic showers may be differentiated from EM showers by their profiles: hadronic showers typically develop over a longer time, and the particle multiplicity as a function of depth is different [73].

The electromagnetic calorimeter is the closer of the two calorimeters to the beamline. Throughout the entire rapidity range over which EM calorimetry is possible, ATLAS employs liquid argon technology – an accordion-type calorimeter in the barrel region $|\eta| < 1.475$, and a specialised design in the forward regions ($1.375 < |\eta| < 3.2$). Each of these three subsections is housed in its own calorimeter. The main section comprises a lead-liquid argon (or LAr) sampling detector, which contains accordion-shaped electrodes and lead absorber plates. As with other detector sections, it is

subdivided into a central barrel and two endcaps. The accordion design ensures that coverage extends throughout the azimuthal plane.

The depth of an electromagnetic calorimeter (and, by extension, the depth of an electromagnetic shower) is typically expresed in terms of the radiation length, denoted X_0; this is the distance of shower travel inside the calorimeter after which the mean particle energy has decreased by a factor 1/e. Notably, the radiation length depends upon the material from which the detector is made (because the rate of shower energy attenuation due to Bremsstrahlung varies between different materials).

The ATLAS EM calorimeter covers $22X_0$ in the barrel region and $24X_0$ in the endcaps; it is divided into three sampling regions of differing granularity, running radially outwards from finest to coarsest. The loss of energy by particles as they traverse the detector upstream of the EM calorimeter is corrected for by the use of a presampler detector at $|\eta| < 1.8$ [74]. The energy resolution of the EM calorimeter is given by the formula

$$\frac{\sigma_E}{E} = \frac{(10.1 \pm 0.4\,\%)(\mathrm{GeV})^{\frac{1}{2}}}{\sqrt{E}} \oplus (0.2 \pm 0.1)\,\%\,, \tag{4.7}$$

where the first (stochastic) term represents the noise and the second constant term represents the effect of non-uniformity and miscalibration.

The forward calorimeter or FCAL comprises copper rods parallel to the beam axis, contained in an outer tube with a liquid argon gap of 250 μm (small so as to prevent ion buildup from the high ionising radiation flux); these tubes are inserted into a copper matrix, and output is read directly from the rods [75].

Directly beyond the EM calorimeter is the hadronic calorimeter – a sampling calorimeter which uses scintillating tiles as the active material and iron as absorber material. It is subdivided into a barrel (covering $|\eta| \leq 1.0$ and sometimes referred to in isolation as the tile calorimeter) and two extended barrels, which cover $0.8 \leq |\eta| \leq 1.7$; both of these regions are divided azimuthally into 64 modules. The outputs of these barrel tiles are read out from either end by wavelength shifting fibres into two separate photomultiplier tubes, which in effect digitises the deposited energy. The extended barrels, in contrast to the barrel, use a copper absorber and liquid argon as the active material; they each comprise two longitudinal segments with four layers.

The hadronic calorimeter is the coarser of the two ATLAS calorimeters, with a typical granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, constrasting with the finer granularity

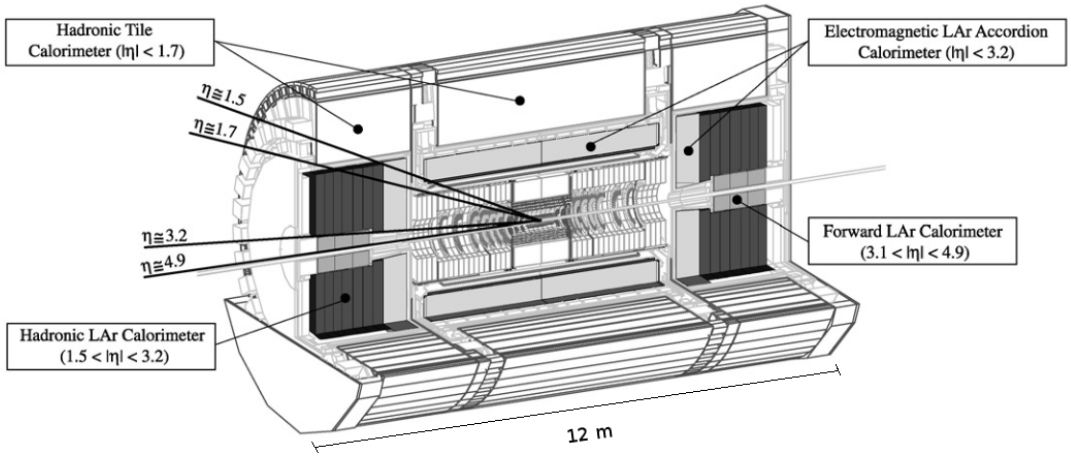


Figure 4.3: Diagram of the ATLAS calorimeters. Adapted from Ref. [75].

of the electromagnetic calorimeter ($\Delta\eta \times \Delta\phi = 0.050 \times 0.025$ in its coarsest layer). Nonetheless, it is capable of identifying the shower geometries associated to certain decay processes, which plays an important role in particle identification (for instance, the shape differences between τ jets and hadronic jets).

For the tile calorimeter, the energy resolution has been measured using beams of pions, and it is described by

$$\frac{\sigma_E}{E} = \frac{56.4 \pm 0.4\,\%(\mathrm{GeV})^{\frac{1}{2}}}{\sqrt{E}} \oplus (5.5 \pm 0.1)\,. \tag{4.8}$$

Notably, this resolution is significantly lower than that of the EM calorimeter, as would be expected from its coarser structure. The energy resolution of the hadronic endcap is

$$\frac{\sigma_E}{E} = \frac{70.6 \pm 1.5\,\%(\mathrm{GeV})^{\frac{1}{2}}}{\sqrt{E}} \oplus (5.8 \pm 0.2)\,. \tag{4.9}$$

4.2.3 The Muon System

The ATLAS muon spectrometer, shown in Figure 4.4, is designed with the intention of performing accurate measurements of muon properties at the terascale. It is immersed in a powerful 3.9 T-peak magnetic field from the superconducting air-core toroidal magnets, which curves the trajectories of the muons. The magnetic field is able to provide 1.5 to 5.5 T m of bending power in the barrel, and 1.0 to 7.5 T m in the endcaps.

High-transverse-momentum muons provide an important signature for many physics

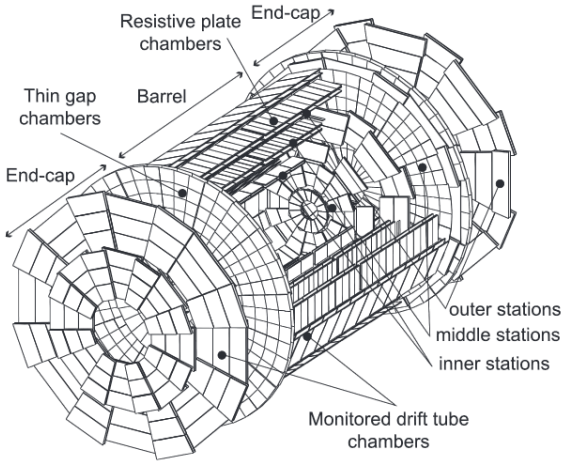


Figure 4.4: Diagram of the ATLAS muon spectrometer, with major subsystems indicated. Taken from Ref. [76].

processes of interest to ATLAS physicists – such as Higgs decays to W^+W^- and ZZ pairs, which subsequently decay to leptons – so careful design of the muon trigger and tracking systems was necessary. The specific design criterion was to attain 2–4 % resolution performance for muons with p_{T} in the range of 10–200 GeV and 10 % for 1 TeV muons; testbeam experiments have shown that this objective has been attained [77, 78].

The muon detector uses two distinct trigger detector technologies: the RPC (Resistive Plate Chambers) in the barrel region, and the TGC (Thin Gap Chambers) in the endcaps. In both regions of the detector, the MDT (Monitored Drift Tube) chambers give precision tracking and momentum measurement, but in the innermost barrel regions, a separate technology is used – CSC (Cathode Strip Chambers). The total number of readout channels is approximately one million.

One important design criterion of the muon spectrometer is its alignment, which aims for a precision of better than 10 µm in order to ensure that momenta can be measured accurately. Thus the muon system is equipped with thousands of sensors for making precision alignment measurements of the device. Further careful calibration is necessary in order to measure and correct for perturbations in the magnetic field caused by other ATLAS components. The effect of the bending power of the spectrometer's magnetic field is quantified in terms of a sagitta, defined as the distance between the centre of the reconstructed track in the middle chamber from the straight line connecting the centres of the reconstructed centres in the outer chambers (*i.e.* the maximum linear displacement of the circular trajectory from a straight line) [79].

The nominal reconstruction accuracy of the muon spectrometer is

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 10\,\%\,. \tag{4.10}$$

Note that this is valid at a reference muon energy of 1 TeV (the precision of the measurement decreases with energy; for energies in the range 10–200 GeV, the precision is 2–3 %).

4.2.4 The Magnet system

ATLAS features three magnet systems, one of which is a solenoid and three of which are toroids; their main purpose is to deflect the trajectories of charged particles via the Lorentz force. This is valuable since the radius of curvature of such a particle depends directly on its charge-to-mass ratio, and it forms an important component of measurements taken in the inner detector and muon spectrometer.

The solenoid is the central magnet, which immerses the inner detector in a 2 T field; this field runs parallel to the beamline. The muon spectrometer's barrel region is immersed in a magnetic field from the barrel toroid (0.5 T) and each of the muon spectrometer's endcaps has its own 1 T toroidal magnet. In total, the fields produced by these four magnets contain some 1.6 GJ of electromagnetic energy, and they permeate a volume of 12 000 m^3.

4.2.5 The Forward Detectors

There are further calorimeters embedded in the end-cap cryostat – the forward calorimeters, or FCAL, which are intended to allow energy measurements in the highest-pseudorapidity regions possible (nominally $3.1 \leq |\eta| \leq 4.9$). The FCAL has a modular construction with three subsystems: a copper and liquid argon module for EM calorimetry, followed by a liquid argon module and a tungsten module, for hadronic calorimetry. Its granularity is very coarse, as can be seen from its resolution function:

$$\frac{\sigma_E}{E} = \frac{100\,\%\,\left(\mathrm{GeV}^{\frac{1}{2}}\right)}{\sqrt{E}} \oplus 10\,\%\,. \tag{4.11}$$

Also in the forward region, far beyond the interaction point, are systems such as LUCID (LUminosity measurement using Cherenkov Integrating Detector, 17 m in

both directions from the interaction point) and ALFA (Absolute Luminosity for ATLAS, $\pm 240\,\mathrm{m}$ from the interaction point), which measure the total luminosity delivered by the LHC to ATLAS.

For heavy-ion running, there is an additional system called ZDC (Zero-Degree Calorimeter), which uses layers of tungesten and quartz rods to determine the centrality of heavy-ion collisions. Because of its distance from the interaction point ($\pm 140\,\mathrm{m}$) it is capable of measuring at high pseudorapidities ($|\eta| \geq 8.2$).

4.2.6 Triggers and Data Acquisition (TDAQ)

Owing to the high collision rates arising from the LHC's unprecedented luminosity, physics data are generated inside ATLAS at a rapid rate, on the order of petabytes per second. It is not feasible for the experiment to record more than an extremely small fraction of this data, and therefore great care must be taken to ensure that the subset of available data which is recorded is as likely as possible to be relevant to the scientific objectives of the experiment. This is the motivation for triggering.

TDAQ is divided into subsystems to parallel the structure of the detector, with three fundamental levels of triggering: Level 1, High Level Trigger (HLT) and Event Filter (these latter two are collectively known as "Level 2"). At each level, the rate of recorded events is largely reduced, by filtering out events that do not meet a strict set of criteria. At Level 1, these decisions are made rapidly by fast electronics in the detector, drawing on information from the various subsystems, and no information from rejected events is permanently retained.

The Level 1 trigger is designed to look for evidence of high-p_{T} leptons, photons or jets, as well as considering missing and total transverse energy, to decide whether a given event is likely to contain new physics. This decision is made quickly (typically on a timeframe of less than $25\,\mu\mathrm{s}$) and so only low-granularity information (for instance, from the calorimeters) is considered. This rapid decision time means that the Level 1 trigger has a very high rate, of the order of 75 kHz. The central trigger processor then implements the trigger menus and prescaling that have beens set. If the event passes Level 1, the event is transferred directly to Data Aquisition (DAQ). The purpose of DAQ is to allow the different levels of triggering to exchange information with one another and to access data relevant to triggering directly from the relevant subsystems of the detector.

The final role of the Level 1 trigger is to define regions of interest (RoI) for the Level

2 trigger. The HLT then processes the event in approximately 40 ms, focussing on the regions of interest that were identified at Level 1. At this level, the trigger rate is approximately 2.0–3.5 kHz. Events which pass the HLT then move to the Event Filter, an offline trigger which operates on a timescale of several seconds. At this stage, the trigger rate (and thus the final event recording rate) is of the order of a few hundred events per second; each event is of the order of one megabyte of data. Events which pass triggering completely are stored permanently by the DAQ at the CERN Computing Centre (CCC).

At the highest level, the operation and monitoring of all subsystems of ATLAS are coordinated by the Detector Control System (DCS), which is also responsible for interfacing between ATLAS and other CERN facilities, such as the LHC accelerator.

4.3 Summary

ATLAS is a complex machine consisting of several subsystems which must work together with high precision in order to generate useful data for physics analyses. It has been designed to provide accurate measurements of physical processes at much higher energies and with much higher luminosities than could be attained at previous experiments, and with a view to studying as wide a variety of high-energy physics phenomena as possible, particularly new physics and new particles at the terascale.

Chapter 5

MadGraph Simulation Studies

"Necessità 'l c'induce, e non diletto."

— Dante, *Divina Commedia*

5.1 Monte Carlo Generators

In order to make quantitative comparisons between theoretical predictions (*i.e.* phenomenology) and observed data, it is necessary to generate sets of simulated events (*Monte Carlo* or MC) which correspond to the predictions of different models. Typically, Monte Carlo generators will simulate physics at the hard-scattering level, delivering their output in a vector format – the Les Houches Event Format (LHEF). In these files, each event is represented simply as a set of particle four-vectors; these collections of four-vectors can then be used as inputs to a separate piece of software which carries out showering, *i.e.* the simulation of the decay and hadronisation of the final state products of the hard scattering process, thus in turn simulating the hadron shower that the detector will actually observe. After this point, one can simulate the downstream experimental analysis using further software; in these studies, this stage was carried out in Rivet [80]. Rivet is a software infrastructure for the cross-validation of Monte Carlo event generators and generator tunings; it is implemented as a library of C++ classes, and its guiding design principle is to ensure portability and flexibility.

Crucially, decisions about how various theoretical models can be implemented within a computational framework are rarely unique, often leading to differences in the predictions of different MC software. In order to understand this effect, it is necessary to compare directly the predictions of different MC generators and to understand

what variables and parameters are most relevant to any observed discrepancy. It is important that MC be robust in the sense that its predictions about observations should not vary significantly as a function of the assumptions implicit in the modelling of the Monte Carlo software; this would clearly be unphysical.

Such potential problems arise because MC is necessarily an approximation to what one would expect to observe from theory – at every stage, an analysis is constrained by computational capabilities, and hence physicists must make a decision about where the optimal tradeoff between computing time and accuracy lies. The approximations made in different generators are not exactly the same, and it is difficult to predict in advance what the consequences of these differences could be for the distributions of observables.

5.2 $t\bar{t}b\bar{b}$ Modelling

In the search for the Higgs boson in the $t\bar{t}H, H \to b\bar{b}$ channel, it is of great importance to understand the modelling of irreducible backgrounds such as $t\bar{t}b\bar{b}$, which is an important contribution to the overall $t\bar{t}$+jets background. To this end, studies were undertaken with the aim of comparing the implementation of this process in the MadGraph 5 generator [81] with that of Alpgen [82]. This was of particular interest due to the potential limitations of Alpgen modelling in certain aspects of ongoing $t\bar{t}H$ studies, particularly relating to the inability to study the decays of top quarks within Alpgen and the necessity of producing separate exclusive samples for $t\bar{t}$ plus light and heavy flavours, with no appropriate parton-jet matching available. Matching is of key importance in the $t\bar{t}H, H \to b\bar{b}$ analysis because of the key role played by heavy-flavour jets – in order to select an event for the analysis, it is essential that the jets in the event are well-matched to partons.

These limitations meant that Alpgen was unable to describe the full $t\bar{t}$+jets background in a satisfactory manner, leading in turn to a desire to continue these studies with MadGraph Monte Carlo instead. The key advantage of MadGraph (a multileg generator, like Alpgen) is that it allows the production of fully inclusive $t\bar{t}$ samples, *i.e.* samples containing $t\bar{t}$ with both light- and heavy-flavour jets, with which matching to all parton flavours is possible.

The reference Alpgen MC was officially-produced ATLAS MC12, available for running on the grid. MadGraph was run locally, using a variety of parameters and internal settings, with the aim of reproducing the event-shape distributions predicted

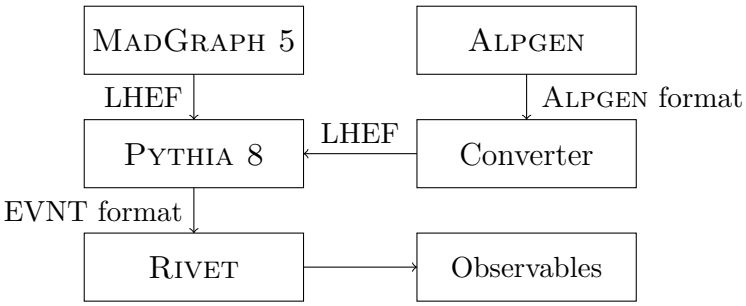


Figure 5.1: Diagram showing the workflow of the MadGraph and Alpgen sample generation and analysis. The output format of each stage is indicated.

by Alpgen. Both MadGraph and Alpgen samples were produced at a centre-of-momentum energy of 8 TeV. The main variable used for making comparisons was $H_{\mathrm{T}}^{\mathrm{jets}}$, *i.e.* the scalar sum of the p_{T} of all jets in the event; this was also the variable used in the $t\bar{t}H, H \to b\bar{b}$ analysis fit within the background-dominated regions. It was also advantageous to consider the p_{T} of the leading jet, as well as the jet multiplicity. This allowed the comparison of changes in the shape of the $H_{\mathrm{T}}^{\mathrm{jets}}$ spectrum with changes in the p_{T} spectra of the individual jets.

In order to produce the final event-shape distributions that were used for making the comparisons, it was necessary to *shower* the LHEF files from MadGraph using Pythia 8 [83] within the ATLAS Athena environment of offline data analysis and simulation tools [84] – originially developed from the LHCb Gaudi framework of data processing applications [85], Athena now enjoys wide use across the ATLAS collaboration. This procedure simulated the evolution of parton showers through the ATLAS detector, taking the hard-scattering partons from the LHEF files as the starting point. The output of the showering was an EVNT file (or "poolfile") which could then be used as input to a Rivet analysis, which then calculated and plotted the observables. The raw Alpgen events were converted from the Alpgen output format to LHEF, then showered and analysed in exactly the same manner. The full workflow is illustrated in Figure 5.1.

5.3 Rivet Analysis

To compare the Alpgen and MadGraph samples, the showered events are analysed using the Rivet analysis framework [80]. Within Rivet, a selection is applied and all relevant variables are calculated. The Rivet analysis code used in these simulation studies implemented the following selection criteria, which are intended to

make the selection as similar to the standard ATLAS $t\bar{t}H$ event selection as possible.

Events are selected based on the semileptonic decay of the $t\bar{t}$ pair. Events are required to have one isolated lepton, which may be either an electron or a muon. For electrons, the requirements $p_{\mathrm{T}} \geq 25\,\mathrm{GeV}$ and $|\eta| \leq 2.47$ are applied; for muons the equivalent requirements are $p_{\mathrm{T}} \geq 25\,\mathrm{GeV}$ and $|\eta| \leq 2.5$. In both cases, the requirements $E_{\mathrm{T}}^{\mathrm{miss}} \geq 20\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{\mathrm{miss}} + M_{\mathrm{T}_W} \geq 60\,\mathrm{GeV}$ are also applied, where $E_{\mathrm{T}}^{\mathrm{miss}}$ denotes the total missing transverse energy in the final state and M_{T_W} denotes the transeverse mass of the W boson.

All jets are required to have $p_{\mathrm{T}} \geq 25\,\mathrm{GeV}$ with a centrality requirement of $|\eta| \leq 2.5$. Jet reconstruction is implemented via the FastJet package, using the anti-k_t algorithm [86] with cone parameter $\Delta R = 0.4$. b-tagging is not implemented, so all b-jets are truth b-jets.

Reconstruction of the $t\bar{t}$ system is done by separately reconstructing the hadronic and leptonic W bosons, then the tops, from their respective decay products. The hadronic W is reconstructed by considering pairs of light jets (indices i, j), calculating the invariant mass m_{ij} of the dijet system in each case, and computing the minimum χ^2 of any jet pair according to

$$\chi^2_{ij} = \frac{1}{\sigma^2_{MW}} \left| M^2_W - m^2_{ij} \right|, \tag{5.1}$$

where M_W is the mass of the W and σ_{MW} denotes the Breit-Wigner width of the W-mass peak. The leptonic W is reconstructed by computing a χ^2 according to

$$\chi^2 = \frac{1}{\sigma^2_{MW}} \left| M^2_W - m^2_l \right|, \tag{5.2}$$

where m_l denotes the mass of a lepton in the event. There is an additional requirement imposed on the reconstructed leptonic W, such that

$$\chi^2 \left(p_{||}^{\mathrm{lep}} \right)^2 - \left(p_{\mathrm{T}}^{\mathrm{lep}} p_{||}^{\mathrm{lep}} p_{\mathrm{T}}^{\mathrm{miss}} \right)^2 > 0, \tag{5.3}$$

where p_{T} denotes either the missing transverse momentum or the transverse momentum of the hard lepton as appropriate, and $p_{||}^{\mathrm{lep}}$ denotes the longitudinal momentum of the hard lepton.

For the reconstruction of the tops, at least two b-jets are required. Chi-square fits are carried out for both the hadronic and leptonic tops, using the formulæ

$$\chi^2 = \frac{1}{\sigma_{MT}^2} \left| M_{\mathrm{T}}^2 - (p_b + m_{W_{\mathrm{had}}})^2 \right| \tag{5.4}$$

and

$$\chi^2 = \frac{1}{\sigma_{MT}^2} \left| M_{\mathrm{T}}^2 - (p_b + m_{W_{\mathrm{lep}}})^2 \right| \tag{5.5}$$

respectively, where σ_{MT} denotes the Breit-Wigner width of the top quark, M_{T} is the mass of the top quark, p_b is the absolute value of the b-quark momentum, and $m_{W_{\mathrm{had}}}$ and $m_{W_{\mathrm{lep}}}$ are the reconstructed masses of the hadronic and leptonic W bosons respectively.

5.4 Comparisons

5.4.1 Default Settings of Alpgen and MadGraph

Initially, MadGraph was used to produce $t\bar{t}b\bar{b}$ events with the default MadGraph run and parameter settings, to understand how these compared with the Alpgen defaults. These comparisons formed the baseline for the subsequent comparisons; the differences between the MadGraph and Alpgen distributions for each adjsutment to the MadGraph tuning were to be evaluated with reference to this first run.

The MadGraph sample included lepton-plus-jets and dilepton events, as well as allowing taus. The comparison of the $H_{\mathrm{T}}^{\mathrm{jets}}$ (*i.e.* the scalar sum of the p_{T} for all jets in the event), p_{T}^{j1} (*i.e.* the p_{T} of the highest-p_{T} jet), and n_j (*i.e.* the jet multiplicity) spectra for these two samples can be seen in Figure 5.2 (a, b, c).

Relevant parameters that could be adjusted within MadGraph included the cross section to normalise to, the masses of the top, bottom and charm quarks, the width of the top (in MadGraph only, since top decays cannot be simulated in Alpgen). Moreover, it is possible within MadGraph to specify the scale function to be used in the matrix element calculation. The choice of software used in showering the LHEF files is in principle independent of the generator (since each generator can be used to prepare events in the standard Les Houches format, the processing downstream of this stage is generator-independent), and it is also possible to choose which PDF set is used in the showering.

Table 5.1 shows those parameters whose values were kept the same across the com-

Param/Setting	Value
σ	1.4353 pb
m_t	172.5 GeV
m_b	4.7 GeV
m_c	0 GeV

Table 5.1: Generator settings used across all Alpgen/MadGraph comparisons.

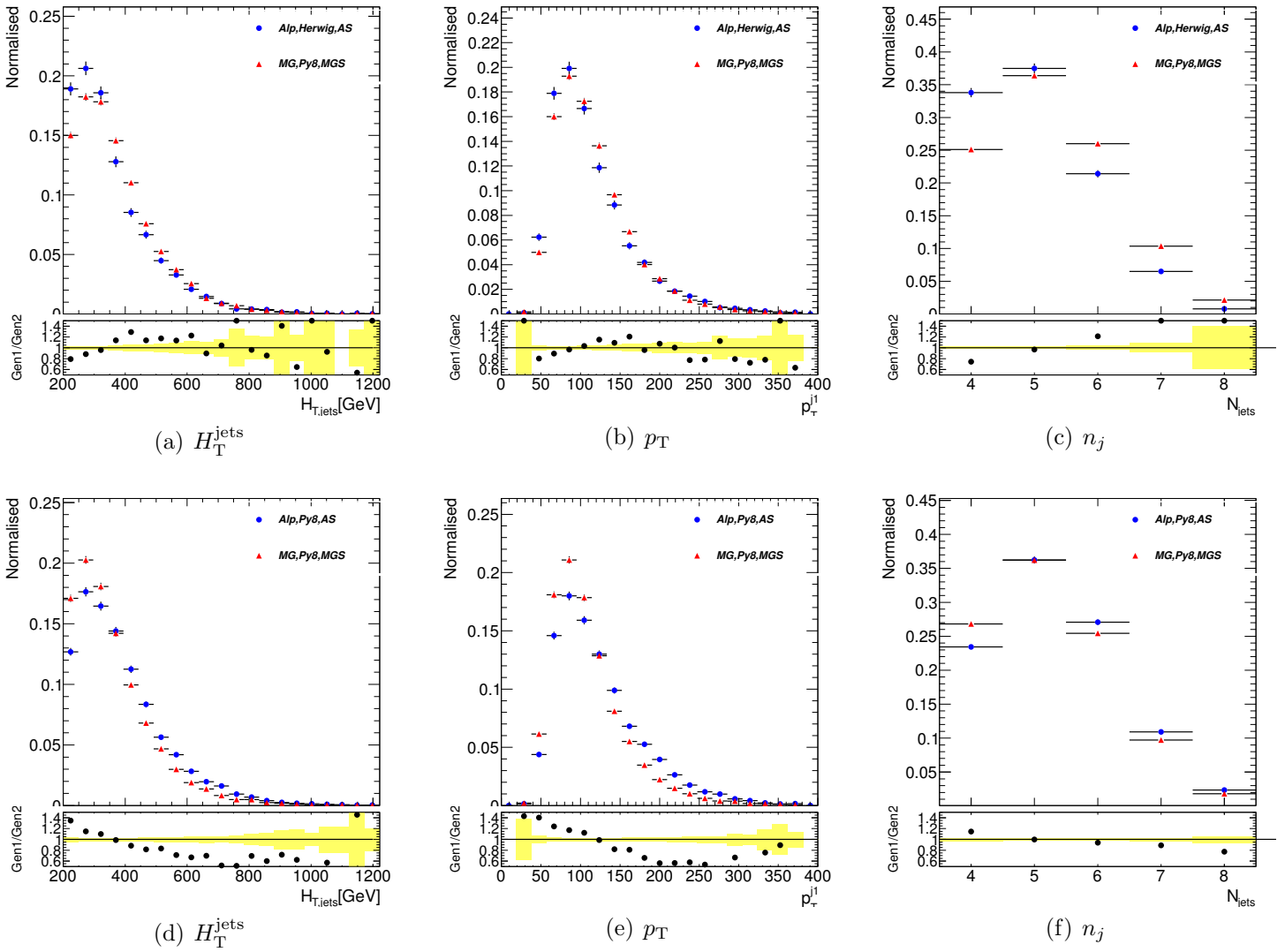
parisons; tables 5.2 and 5.3 show the relevant settings that were used to prepare the Alpgen and MadGraph samples used in each comparison.

	Comparison 1 – Default		Comparison 2 – Same Showering		Comparison 3 – Same Scale	
Param/Setting	Gen. 1 (blue dot)	Gen. 2 (red triangle)	Gen. 1 (blue dot)	Gen. 2 (red triangle)	Gen. 1 (blue dot)	Gen. 2 (red triangle)
Generator	Alpgen	MadGraph 5	Alpgen	MadGraph 5	Alpgen	MadGraph 5
No. events	1743500	150000	325076	150000	1743500	150000
asrwgtflavor	(MG Only)	5	(MG Only)	5	(MG Only)	5
maxjetflavor	(MG Only)	4	(MG Only)	4	(MG Only)	4
Top width (GeV)	0	1.4915	0	1.4915	0	1.4915
Top decays	N	Y	N	Y	N	Y
Cuts on top decay products	N	Y	N	Y	N	Y
PDF in Generator	CTEQ6L1	CTEQ6L1	CTEQ6L1	CTEQ6L1	CTEQ6L1	CTEQ6L1
Scale	Alpgen	MadGraph	Alpgen	MadGraph	Alpgen	Alpgen
Showering	Herwig	Pythia 8	Pythia 8	Pythia 8	Herwig	Pythia 8
PDF in showering	CTEQ6L1	MSTW2008LO	MSTW2008LO	MSTW2008LO	CTEQ6L1	MSTW2008LO

Table 5.2: A table of the generator settings used in the first three Alpgen/MadGraph comparisons. The flags "asrwgtflavor" and "maxjetflavor" are internal generator tuning parameters exclusive to MadGraph (denoted by "MG only"), and their values are reproduced here for the sake of completeness. The different scales are as defined in Section 5.4.3. The generators used in showering are as explained in Section 5.4.3, and the PDF sets are as explained in Section 5.4.5.

	Comparison 4 – Same Showering & Scale		Comparison 5 – Undecayed $t\bar{t}$ System		Comparison 6 – Optimal Agreement	
Param/Setting	Gen. 1 (blue dot)	Gen. 2 (red triangle)	Gen. 1 (blue dot)	Gen. 2 (red triangle)	Gen. 1 (blue dot)	Gen. 2 (red triangle)
Generator	Alpgen	MadGraph 5	Alpgen	MadGraph 5	Alpgen	MadGraph 5
No. events	325076	150000	325076	150000	1743500	150000
asrwgtflavor	(MG Only)	5	(MG Only)	5	(MG Only)	5
maxjetflavor	(MG Only)	4	(MG Only)	4	(MG Only)	4
Top width (GeV)	0	1.4915	0	0	0	0
Top decays	N	Y	N	N	N	N
Cuts on top decay products	N	N	N	Y	N	N
PDF in Generator	CTEQ6L1	CTEQ6L1	CTEQ6L1	CTEQ6L1	CTEQ6L1	CTEQ6L1
Scale	Alpgen	Alpgen	Alpgen	MadGraph	Alpgen	Alpgen
Showering	Pythia 8	Pythia 8	Pythia 8	Pythia 8	Pythia 8	Pythia 8
PDF in showering	MSTW2008LO	MSTW2008LO	MSTW2008LO	MSTW2008LO	CTEQ6L1	CTEQ6L1

Table 5.3: A table of the generator settings used in the latter three Alpgen/MadGraph comparisons. The flags "asrwgtflavor" and "maxjetflavor" are internal generator tuning parameters exclusive to MadGraph (denoted by "MG only"), and their values are reproduced here for the sake of completeness. The different scales are as defined in Section 5.4.3. The generators used in showering are as explained in Section 5.4.3, and the PDF sets are as explained in Section 5.4.5.



(a) $H_{\mathrm{T}}^{\mathrm{jets}}$ (b) p_{T} (c) n_j

(d) $H_{\mathrm{T}}^{\mathrm{jets}}$ (e) p_{T} (f) n_j

Figure 5.2: Alpgen/MadGraph comparisons using default settings ("Comparison 1", upper row) and consistent Pythia 8 showering ("Comparison 2", lower row). "Gen1" and "Gen2" denote the first and second generators as shown in Table 5.2.

As can be seen immediately from the ratio plots in Fig. 5.2 (a, b, c), the agreement between the two generators is poor, and there is a very significant shape difference between the two H_{T} spectra, which can also be seen in the p_{T} spectra of the leading jet. Moreover, the jet multiplicity distributions suggest that Pythia 8 produces harder showers than Herwig.

These observed differences prompted a careful consideration of possible sources of discrepancy between the two generators, as well as differences in the downstream treatment of the samples. In particular, the Alpgen sample (on the grid) had been showered in Herwig, whereas the MadGraph had been showered in Pythia 8, and the default renormalization and factorization scales differed significantly between the two generators. These sources of discrepancy were considered separately.

5.4.2 Showering

The choice of generator for showering the Monte Carlo is important, since different generators produce showers of varying hardness. Shower hardness could potentially have a large effect on the $H_{\mathrm{T}}^{\mathrm{jets}}$ spectra, so the optimal solution is to make comparisons only between samples that have been showered in the same generator. Since showering job options interfacing MadGraph to Herwig were unavailable at the time of these studies, it was decided to convert the raw Alpgen output (in Alpgen's native format) to the Les Houches Event Format, and then to shower it in Pythia 8, following the same procedure that was used with the MadGraph output. This enabled a fairer comparison of the two sets of Monte Carlo samples, which can be seen in Fig. 5.2 (d, e, f). As can be seen from the remaining large differences in the spectra shapes, changing the showering alone did not greatly improve the overall Alpgen/MadGraph agreement.

5.4.3 Scales

The matrix element calculations performed by generators such as Alpgen and MadGraph take place at a certain energy scale, which can be set by hand in their run cards. This scale can be set to a fixed value (*e.g.* the mass of the Z boson) or it can be dynamic.

A dynamic scale takes the form of a relationship between the renormalisation scale μ_R and the factorisation scale μ_F. In the simulation studies described here, there are two different dynamic scales of interest – the MadGraph default scale

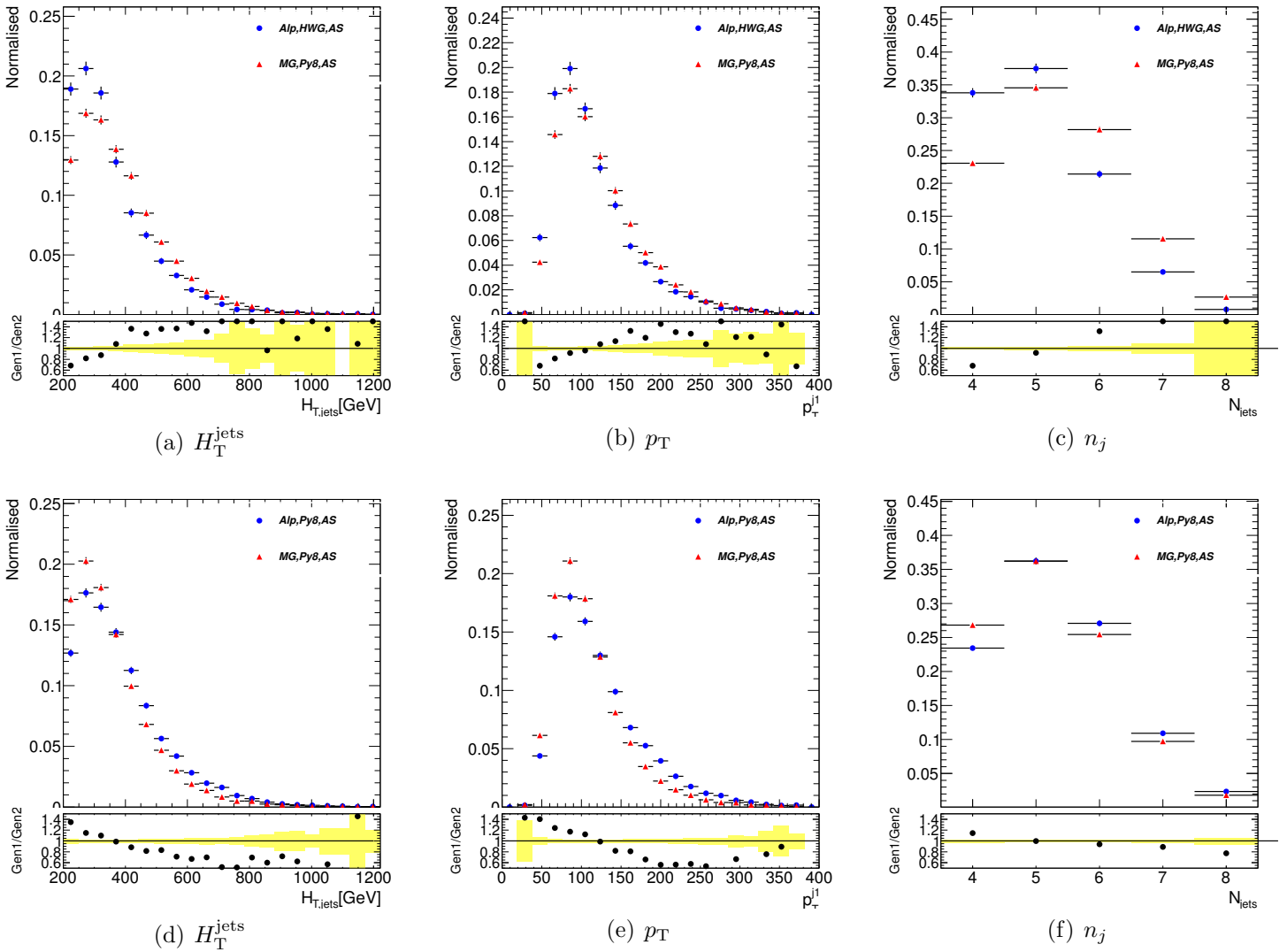
$$\mu_{\mathrm{F}}^2 = \mu_{\mathrm{R}}^2 = Q^2 \left(M_{\mathrm{FS}}^2 + \Sigma p_{\mathrm{T}}^2 \right), \tag{5.6}$$

where Q^2 is the scale factor, M_{FS} is the maximum mass of a final-state particle, and the sum is over jets and visible massless particles [87], and the Alpgen default scale

$$\mu_{\mathrm{F}}^2 = \mu_{\mathrm{R}}^2 = \Sigma m_{\mathrm{T}}^2 = \Sigma \left(m^2 + p_{\mathrm{T}}^2 \right), \tag{5.7}$$

where the sum is over jets and visible massless particles.

By editing the MadGraph configuration files it was possible to implement the Alpgen scale in MadGraph, to eliminate scale difference as a source of discrepancy between the generators. An exercise to validate the implementation of the Alpgen scale in MadGraph is detailed in Appendix A.



(a) $H_{\mathrm{T}}^{\mathrm{jets}}$ (b) p_{T} (c) n_j

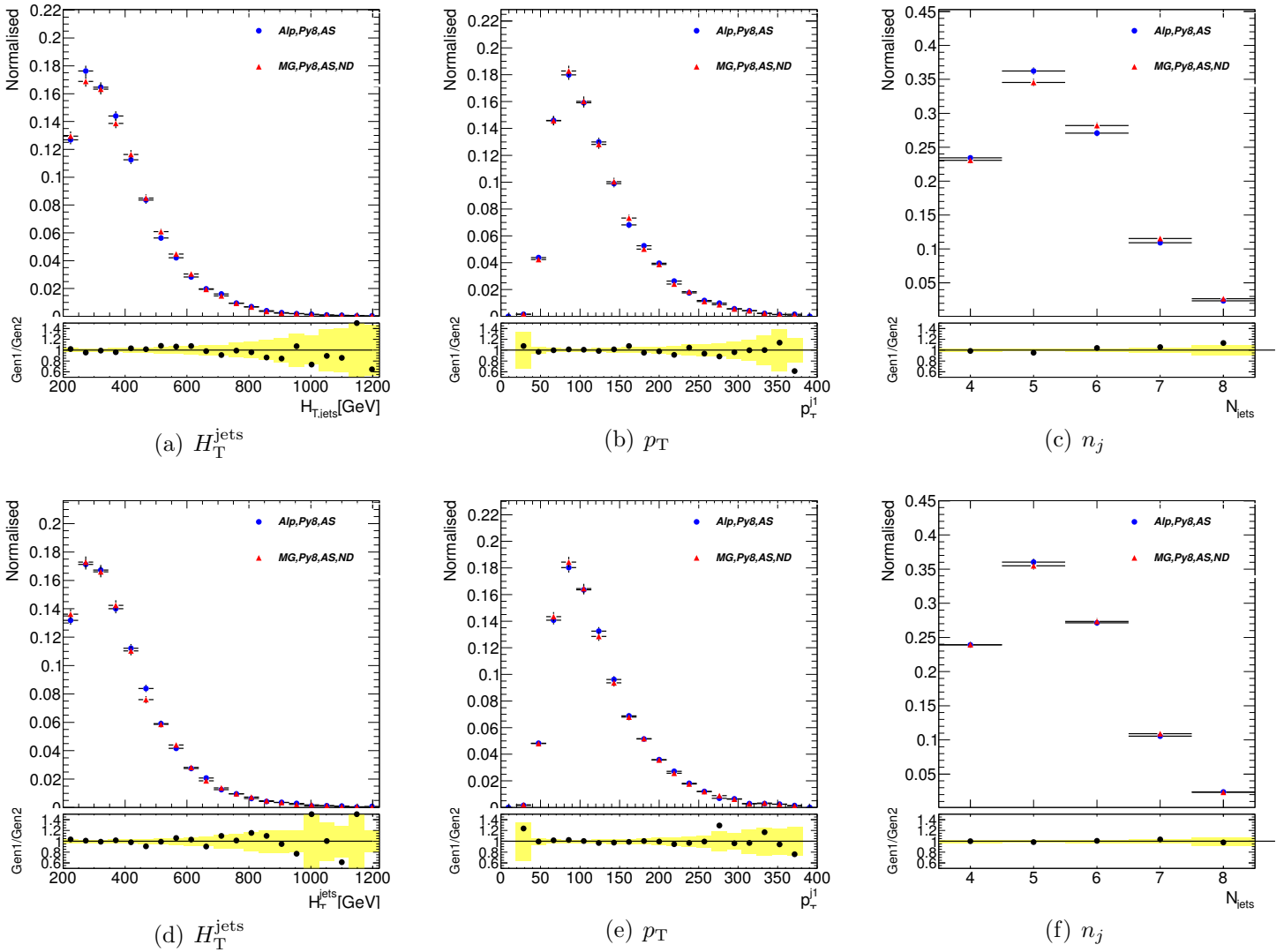
(d) $H_{\mathrm{T}}^{\mathrm{jets}}$ (e) p_{T} (f) n_j

Figure 5.3: Alpgen/MadGraph comparisons with same scale ("Comparison 3", upper row), and consistent Pythia 8 showering and scale choice ("Comparison 4", lower row). "Gen1" and "Gen2" denote the first and second generators as shown in Tables 5.2 and 5.3.

As can be seen in Figure 5.3 (a, b, c), implementing the Alpgen scale in MadGraph did not produce a noticeably greater agreement between the generators. It therefore remained to combine these two steps and compare the MadGraph sample at the Alpgen scale against the Alpgen sample with Pythia 8 showering. However, the comparison with consistent Pythia 8 showering as showing in Figure 5.3 (d, e, f) shows an agreement which is still not optimal. This was suspected to be caused by asymmetric treatment of top decays in the two samples, necessitating the regeneration of the MadGraph sample without top decays.

5.4.4 Treatment of top decays

A further important source of disagreement between the samples lay in the treatment of top quarks and their decays – specifically, Alpgen is unable to decay the tops in $t\bar{t}b\bar{b}$ events, whereas the earlier MadGraph samples all featured explicitly decayed tops. This was problematic because the decay would alter the kinematics of the event and the scale calculation (because the dynamic scale depends upon the masses of the final-state particles from the hard-scattering calculation, which are the top decay products in the case of decayed tops, hence less massive than the top quarks themselves). Moreover, MadGraph by default imposed phase-space cuts on the top decay products, further differentiating it from Alpgen. It was therefore of great importance to ensure that the treatment of tops was thoroughly understood and identical between the two generators. As a validation exercise, an additional MadGraph sample was prepared with top decays enabled, but without the default cuts imposed on the decay products. This would ensure that the same region of phase space was being examined in both cases. In Figure 5.4 it can be seen that, in line with expectation, disabling top decays completely, while also incorporating the improvements in scale and showering, greatly improved the observed agreement between the two generators.



(a) $H_{\mathrm{T}}^{\mathrm{jets}}$ (b) p_{T} (c) n_j

(d) $H_{\mathrm{T}}^{\mathrm{jets}}$ (e) p_{T} (f) n_j

Figure 5.4: Alpgen/MadGraph comparisons with undecayed $t\bar{t}$ system ("Comparison 5", upper row), and consistent shower PDF set, Pythia 8 showering, scale choice and $t\bar{t}$ treatment ("Comparison 6", lower row). "Gen1" and "Gen2" denote the first and second generators as shown in Table 5.3.

5.4.5 Reshowering with the CTEQ6L1 PDF Set

A key choice in producing MC for physics processes at hadron colliders is that of the set of parton distribution functions (PDFs) [88] to be used within the MC software. These functions model the distribution of momentum among the partons within hadrons (such as protons), and an appropriate PDF choice is essential for the accurate modelling of the QCD processes which define the kinematics of a proton-proton interaction.

To increase the consistency of generating and processing the samples, therefore, it was desirable to use the CTEQ6L1 PDF set [89] consistently throughout, in generation and showering, and then to repeat the comparison. Previous comparisons had used CTEQ6L1 at the generator level, but the MSTW2008LO PDF set [90, 91] was used at showering. Table 5.3 (Comparison 6) shows the settings used in the production of these samples; the comparisons in the relevant spectra can be seen in Figure 5.4 (d, e, f). By this consistent use of the CTEQ6L1 PDF set throughout the production of the simulated data sets, a notable improvement in agreement between Alpgen and MadGraph is observed.

5.5 Conclusions

After implementing the same scale in both Alpgen and MadGraph, and after carrying out all showering consistently in Pythia 8 and decaying the tops at the showering stage rather than at the generator level – *i.e.*, by performing the fairest possible comparison – it was found that the agreement in the distributions of observables between the Monte Carlo datasets produced the two generators was greatly improved, yielding compatible predictions. With MadGraph having passed this validation exercise, it was therefore feasible to choose MadGraph as the generator to be used for the production of the inclusive light- and heavy-flavour $t\bar{t}$ background events for the $t\bar{t}H, H \to b\bar{b}$ analysis. The ATLAS analysis ultimately chose to generate $t\bar{t}$ background events in MadGraph with showering and top decays in Pythia, since this was considered to be the optimal combination given the necessity of adequate matching and the limitations imposed by the availability of Athena job options at the time. However, the combination of the Powheg generator [92] with Pythia showering was found to give optimal data-MC agreement, so an appropriate reweighting procedure was devised to optimise the MadGraph MC for data-MC agreement [93].

Chapter 6

Statistics and Data Analysis

> "Le secret d'ennuyer est celui de tout dire."
>
> — Voltaire, *Sept Discours en Vers sur l'Homme*

> "There are conditions worse than being unable to see, and that is imagining one sees."
>
> — LRH

6.1 Probability Theory

Statistics is the mathematical discipline which is concerned with the quantitative analysis of data, with the objective of either describing and classifying relevant features of a data set, or of making probabilistic inferences about the patterns underlying the data set. This can allow one to make inductive arguments from a given set of observations to justify hypotheses about the underlying physics.

Central to statistics is the concept of probability; in modern mathematical usage, this is commonly defined by Kolmogorov's axioms. In this formalism, one speaks of data as events drawn from an event space, or sample space – the events may be thought of as subspaces of the event space. The probability of some event E, denoted $P(E)$, is a measure on the event space, which is required to satisfy three axioms to ensure that the probability is a well-defined measure on the event space.

1. Positive semidefiniteness: $P(E)$ is a finite, non-negative real number for all events E in the event space.

2. Unitarity: the probability of the entire event space is one (if measure on the event space is not well-defined, then neither are the probabilities).

3. For any countable set of mutually-exclusive events $E_1, E_2, ...$ the relation $P(E_1 \cup E_2 \cup ...) = \Sigma_{i=1}^{\infty} P(E_i)$ is satisfied. Intuitively, the probability of a union of discrete events is simply the sum of their respective probabilities.

From these axioms one can immediately deduce some other important properties of probabilities: for events satisfying $E_1 \subseteq E_2$, $P(E_1) \leq P(E_2)$ (*i.e.* probabilities are monotonic functions of their event spaces); for all events E the bound $0 \leq P(E) \leq 1$ applies; and the probability of the empty set is $P(\varnothing) = 0$.

An important observation is that the definition of the event space is in a sense arbitrary; any subspace of an event space may equally be considered an event space in its own right, and thus the same axioms and theorems that apply to probabilities measured on the larger event space must apply equally to probabilities measured on the subspaces. Thus one can define relations between probabilities measured on different subspaces (*i.e.* conditional probabilities). Thus one can speak of such probabilities as $P(A|B)$, the conditional probability of event A given that (*i.e.* measured on the event subspace where) event B obtains. Most importantly, it can be proven that conditional probabilities observe the relation

$$\frac{P(A|B)}{P(B|A)} = \frac{P(A)}{P(B)}, \tag{6.1}$$

which specifies that the ratio of such conditional probabilities is equal to the ratio of the absolute or intrinsic probabilities. This is Bayes's Theorem, and it serves as the foundation for Bayesian inference, a form of statistical reasoning widely employed in many scientific fields, including high-energy physics. It can be understood as a description of how to adjust the estimated *subjective* probability of a hypothesis as new evidence becomes available: thus one speaks of the probability of the signal plus background hypothesis, as contrasted with the null (background-only) hypothesis, given that certain observations have been made. On this model, one understands probabilities not as providing objective descriptions of reality, but as quantifying the relative degree of confidence which a scientist has in different hypotheses. Rewriting (6.1) slightly, the intuitive motivation for regarding it as a description of how to revise an estimated probability in light of new evidence may be seen as

$$P(A) = \frac{P(A|B)}{P(B|A)} P(B). \tag{6.2}$$

This quantifies the "effect" on the subjective probability $P(A)$ of new evidence about the conditional probabilities of observing A given B, and B given A.

The Bayesian approach contrasts with frequentism, another standard interpretation of probability theory. According to frequentism, the probability of an event is defined as the limit of its relative frequency in a large number of trials; the underlying assumption is that, in an arbitrarily large number of experimental trials, the relative frequency of a certain observation will converge to its true frequency in the event space. Notably, the frequentist definition does not make use of prior probabilities; this is a key weakness of the Bayesian approach, because accurate estimates of prior probability are often difficult to make in an objective manner [94, 95].

In frequentist statistics one speaks of p-values, which quantify the probability of obtaining a result lying at least as far from the null-hypothesis expectation as a given observation, on the assumption that the null hypothesis is correct. A suitably low p-value can thus be interpreted as statistical evidence that the null hypothesis (for instance, a background-only hypothesis in which the signal that is being tested for is absent) is false, or more precisely that it is to be rejected in favour of some alternative hypothesis, according to which the result is more probable.[(i)] In line with the assumptions underlying frequentism, it is necessary for particle physicists to gather a sufficiently large data set in order to be able to distinguish between the presence and absence of a signal with ample statistical significance. This is referred to as sensitivity to a given signal [96].

The experimental criteria for evidence and discovery are frequentist, in that physicists speak of *evidence* of at least three standard deviations (3σ) from the null-hypothesis expectation, and the normal threshold for a discovery is 5σ. These are measures of how statistically significant an observed effect is with respect to known background processes; they quantify the probability that the observation is a fluctuation within the backgrounds. For the purposes of particle physics, the number of standard deviations of excess (sensitivity, S) is most straightforwardly approximated as

$$S \approx \frac{N_\mathrm{s}}{\sqrt{N_\mathrm{b}}}, \tag{6.3}$$

where N_s denotes the number of signal events and N_b the number of background events. This approximation is valid only when the total number of events N is

[(i)] *Sensu stricto*, frequentism cannot assign probabilities to hypotheses as it considers only the conditional probabilities of observations given different hypotheses.

large; it can be thought of as equivalent to a signal-to-noise ratio, in that the backgrounds are understood as random fluctuations in the signal [97, 98]. A more precise approximation to S, accurate at lower N, is

$$S \approx \sqrt{2N \ln\left(1 + \frac{N_\mathrm{s}}{N_\mathrm{b}}\right) - 2N_\mathrm{s}} \,. \tag{6.4}$$

The frequentist model also gives the motivation for confidence levels in setting exclusions and limits. If a certain signal has been excluded to a 95 % confidence level (CL), that is equivalent to assigning a p-value of 0.05 to the observation given the background-only hypothesis – *i.e.* there is at most a 5 % probability that such an observation would be made under the assumption that the background-only hypothesis is correct (or equivalently, that the signal does not exist), given available information [99].

In the context of high energy physics, the large number of collisions in a detector yields the large number of measurements necessary for statistical significance. Expected probabilities can be calculated from theory (this is the task of phenomenology) and then compared to data in order to extract a result. The frequentist understanding of statistical inference considers each experimental trial (for instance, each particle collision) to be one of a potentially infinite sequence of such trials, each of which is in principle capable of producing a statistically-independent result. Thus, with a sufficiently large quantity of data, one can simply combine all data points and extrapolate to draw conclusions about the underlying physics.

6.2 Distribution Functions

In understanding the properties of data sets, it is important to understand the nature of distribution functions. The most commonly-encountered of these is the *Gaussian* (or *normal*) distribution, which describes the probability that a given observation will deviate by a certain amount from a mean value (typically denoted μ). The Gaussian distribution describes the phenomenon of statistical *noise* or random fluctuations in the data set. The distribution is characterised by its mean μ in addition to a characteristic width parameter, the standard deviation σ (this parameter may also be understood as the square root of the distribution's *variance*, $\mathrm{Var}(X) = \mathrm{E}\left[(X - \mu)^2\right] = \sigma^2$, where E denotes the expectation value). For any Gaussian, a region of a fixed number of standard deviations in width will contain a corresponding proportion of the total mass of probability. By convention,

the area enclosed by the curve is normalised to unity, in keeping with the unitarity axiom of probability theory. The general form of a Gaussian $f(x)$ is specified by the probability density function (pdf)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right). \tag{6.5}$$

And the normalisation criterion – requiring that the total mass of probability contained in the distribution is equal to unity – is simply

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1. \tag{6.6}$$

More generally, the total probability of an observation falling within a certain region (the *mass* of probability contained therein) may be quantified as the integral of the Gaussian function (the probability *density*) over that region. Thus there is the relation

$$p = \int_{-\infty}^{S} f(x)\,\mathrm{d}x, \tag{6.7}$$

where $f(x)$ denotes a normalised Gaussian function with $\mu = 0, \sigma = 1$. This shows that the p-value and the significance S are equivalent pieces of information: either one can be computed directly from the other. The explicit form of this functional dependence is

$$p = 1 - \Phi(S) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{S}{\sqrt{2}}\right)\right], \tag{6.8}$$

for a normal distribution ($\mu = 0, \sigma = 1$) and following the one-sided convention, where Φ denotes the cumulative distribution function associated to the normal distribution.

Table 6.1 shows the relationship between the p-value and the significance obtained via equation (6.7). The value $S \approx 1.644$ is important as it corresponds to the 95 % confidence level threshold ($p = 0.05$) that is typically used for exclusions. Notably, this table follows the one-sided convention as used by ATLAS, according to which the integration runs only from the centre of the Gaussian outwards in one direction (since the Gaussian is symmetric about its mean, following the two-sided convention would simply double the corresponding integral, and hence double the p-value).

Significance (S)	Approximate p-value
1	0.16
1.644	0.05
2	0.023
3	0.0013
4	3.2×10^{-5}
5	2.9×10^{-7}

Table 6.1: Table illustrating the relationship between the significance and the p-values (quoted to two significant figures), following the one-sided Gaussian convention.

It is important to note that the definition of the p-value given here is that of a *local* p-value, *i.e.* it corresponds to the statistical significance of a single experimental trial considered in isolation. In practice, a real experiment will require many such "trials", resulting in a reduction of *global* significance due to the *look elsewhere effect*. The meaning of this can be understood by considering an arbitrary experiment involving n trials – in this case, one would expect an observed p-value of $\frac{1}{n}$ in the absence of signal. For the simplest analyses, involving n statistically-independent trials, the solution is simply to multiply the p-value by n; the number of trials can serve as a crude means of quantifying the look elsewhere effect. Other analyses – including Higgs analyses – can necessitate a more statistically sophisticated approach, such as computing a "trial factor": the ratio of the probability of observing a local excess of events to the probability of observing a similar excess elsewhere in the range of Higgs masses under consideration [100].

6.2.1 Poisson Statistics

The counting statistics of discrete variables are typically governed by the Poisson distribution; the probability of observing n events in an experiment is generally given by the Poisson function

$$P(n, \langle n \rangle) = \frac{\langle n \rangle^n}{n!} \, \mathrm{e}^{-\langle n \rangle} \tag{6.9}$$

where, in the case of no signal observation, $\langle n \rangle = \langle N_\mathrm{b} \rangle$. Where new physics (a signal) is present, $\langle n \rangle = \langle N_\mathrm{b} \rangle + \langle N_\mathrm{s} \rangle$. In Equation (6.4) above, the term $\sqrt{N_\mathrm{b}}$ may be thought of as the Poisson error on the background, *i.e.* a counting uncertainty. It is

notable that, with an infinitely large number of data points, the Poisson distribution converges to a Gaussian.

In real experiments, it is not possible to extract continuous distributions of observables such as p_{T} spectra. For a large set of discrete observations one can instead produce histograms, which with appropriate binning and a suitably large number of entries will reproduce the features of the underlying physical distribution. One can fit a function (such as a Gaussian) to such a histogram and compute a fit parameter – for instance, via a χ^2 test to quantify how closely the observed data fit the expected distribution. In such a way it is possible to compare observations with different hypotheses, by measuring the closeness of the fit to distributions corresponding to different hypotheses.

6.2.2 Likelihood Estimation and Hypothesis Testing

A more nuanced approach than using cuts to define phase space regions is to attempt to calculate the probability, given all available information, that an event results from the signal process; this has the advantage of being a more genuinely multivariate method. Thus, for each event, one computes a single quantity which can be either used as a single-variable discriminant or fitted to the expected signal and background contributions. Typically one begins such an analysis by attempting to measure the probability density associated to the different bins in the input event space. Then one can define a likelihood function proportional to the probability density function associated to the signal, and use this for signal-background discrimination. Typically, a likelihood function takes the form

$$L(\theta|x) = P(x|\theta), \tag{6.10}$$

where θ denotes a set of parameter values (say, some parameters of the hypotheses under consideration) and x denotes a set of observed values. Thus the likelihood of a certain set of parameters θ, given observations x, is defined as the probability of observing x given θ.

Alternatively, one can adopt a more subtle approach by comparing the likelihood function for the signal with that of the background: this is a likelihood ratio (LR). Such a ratio $\Lambda(x)$, for comparing null hypothesis $H_0 : \theta = \theta_0$ and some alternate hypothesis $H_1 : \theta = \theta_1$, can be written as

$$\Lambda(x) = \frac{L(\theta_0|x)}{L(\theta_1|x)}\,. \tag{6.11}$$

The null hypothesis is rejected in favour of the alternate hypothesis when the criterion $\Lambda(x) \leq \eta$ is met, for some η chosen in order to reach a desired level of statistical significance α, defined as $\alpha = P\left(\Lambda(X) \leq \eta|H_0\right)$ [101, p. 120-137].

In the specific case of Higgs searches, H_0 can be thought of as the hypothesis that all observed data are the results solely of higgsless standard model processes; H_1 would then be the hypothesis that a standard model Higgs boson exists (within a certain mass range) and accounts for some contributions to the observed distributions.

Typically it is important to identify parameters which, while not of direct interest, must nonetheless be accounted for and understood owing to their effects on the analysis. These are the so-called *nuisance parameters*; a commonly-encountered example would be the variance of a distribution for which the mean value is the parameter of direct experimental interest. It is often possible to treat the nuisance parameters by partitioning the likelihood function into the respective contributions from the parameters which are to be measured, and from the nuisance parameters. The LR method can also allow one to compute the significance and confidence intervals for the parameters of interest in a manner which takes account of the nuisance parameters [102].

A number of variations on the LR technique exist, such as the log-likelihood method, in which the difference in the natural logarithms of the likelihoods is used as a proxy to the LR (since the logarithm is a monotonically-increasing function with maxima and minima located at the maxima and minima of its input function, it conserves the important information contained in the LR); one of the chief advantages of this approach is that the derivatives of the log-likelihood function are typically easier to compute. Another widely-used likelihood estimation method is that of profile likelihood, whereby the number of free parameters is constrained by considering that some parameters can be expressed as functions of other parameters; in principle this method can be applied to any analysis, and it can often be used to calculate superior confidence intervals when compared to related techniques. However, its utility is somewhat compromised by the fact that the profile likelihood is not a true likelihood function, in that it does not derive directly from a probability density function [103, 104][101, p. 161].

The LR method is often more suitable, since the absolute likelihood of an event coming from the signal is not necessarily the best discriminating variable, due to

unusual event topologies with low absolute likelihoods, or due to unexpected relationships between the signal and background likelihood functions. According to the Neyman-Pearson lemma, a likelihood ratio test is the most statistically-powerful test for a given α, *i.e.* it is the most likely form of statistical test to reject the null hypothesis if it is false [105]. Thus, the LR approach constitutes a very useful tool for reliably measuring p-values from probability density functions. However, there is another important weakness of the method of likelihood estimation: in bins with very few events, the likelihood ratio can be difficult to compute accurately, and the low statistics can give rise to anomalies, such as deceptively large likelihood ratios.

A likelihood function $\mathcal{L}$ for a binned analysis will in general take the form

$$\mathcal{L}(\mu,\theta) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} \exp\left(-(\mu s_j + b_j)\right) \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} \exp\left(-u_k\right), \tag{6.12}$$

where μ is a parameter denoting the strength of the signal (*i.e.* $\mu = 0$ in the background-only case and $\mu = 1$ in the case where the nominal signal is present), θ represents the set of nuisance parameters (*i.e.* those parameters, such as detector and modelling effects, which can account for some of the variation in the observed signal strength, but which are not of immediate experimental interest). s_i and b_i denote the expected contents of bin i in the signal and background distributions respectively (such that $E\left[n_i\right] = \mu s_i + b_i$ is the expected content of bin i for signal strength μ), n_j denotes the content of bin j, and m_k is the content of bin k in the control histogram. $u_k^{m_k}$ is the expectation value of m_k, when its dependence on the nuisance parameters θ is considered.

The likelihood ratio Q, then, is constructed as a ratio of two such likelihood functions, corresponding to the background-only hypothesis (H_0) and the signal-plus-background hypothesis (H_1) respectively. This takes the form

$$Q = \frac{\mathcal{L}(\text{data}|H_1)}{\mathcal{L}(\text{data}|H_0)} = \frac{\mathcal{L}(\text{data}|\mu s + b)}{\mathcal{L}(\text{data}|\hat{\mu} s + b)}\,, \tag{6.13}$$

where $\hat{\mu}$ is the value of μ at which $\mathcal{L}$ attains its maximum. The log-likelihood ratio *test statistic*, $q_\mu = -2\ln Q$, is the value typically used to decide when a hypothesis may be accepted.

Using simulated data corresponding to the signal-plus-background hypothesis, it is possible to construct a probability density function for the test statistic; typically, this procedure is repeated several times in order to account for the variation in the

nuisance parameters. One can then compare this pdf of the test statistic with its observed value (computed using real data) and extract a p-value, $\mathrm{CL_{s+b}}$, which represents the probability that the observed test statistic will be *at most* as compatible with the hypothesis H_1 as the expected test statistic.

6.2.3 Systematic Uncertainties and Nuisance Parameters

In order to examine the effect of a source of systematic uncertainty (or, equivalently, a nuisance parameter) x on a likelihood-based analysis, one typically constructs distributions of the *pull* defined by

$$g = \frac{x - \mu}{\sigma}\,, \tag{6.14}$$

where μ and σ are the mean and standard deviation of the underlying Gaussian distribution, respectively. It is expected *a priori* that the pull of a systematic should itself be distributed according to a normal distribution (*i.e.* $\mu = 0$, $\sigma = 1$); large deviations from this suggest mismodelling of the systematics. In particular, $\sigma > 1$ implies that the effect of the systematic has been underestimated, while $\sigma < 1$ implies that it has been overestimated. Pulls are initially computed using so-called Asimov data (*i.e.* a representative set of expected data which stands in for real data) [104] in order to verify that the Monte Carlo modelling is consistent. Following unblinding and the fit to real data, a second set of pulls may be computed [106].

6.2.4 Discovery Criteria and Limit Setting

Most simply, one speaks of a certain signal (*e.g.* production of the Higgs) as having been excluded when analysis of existing data implies that there is a probability of at least 95 % that all observations can be explained by background processes (*e.g.* Higgsless standard model processes). Thus, the degree to which a hypothesis has been excluded can be quantified as a function of the confidence interval for a given cross section.

However, it is also possible to think of the exclusion relative to the maximum Higgs production cross section which can be excluded by the analysis. One could say, for example, that a given analysis is capable of excluding the Higgs in principle when the expectation value of the excluded cross section is at most the standard model cross section.

In statistical terms one might also say that the overlap of the probability density functions corresponding to the background-only and signal-plus-background hypotheses corresponds to the difficulty in distinguishing the two hypotheses. The quantity that is typically used to define the exclusion of the signal is CL_s, defined as

$$CL_s = \frac{CL_{s+b}}{CL_b}, \tag{6.15}$$

where CL_b is the confidence level at which the background-only hypothesis is excluded and CL_{s+b} is the confidence level at which the signal-plus-background hypothesis is excluded. This definition is adopted as part of the modified frequentist method, according to which one compensates for the possibilty of observing fluctuations in the background distribution resembling the signal, even when the signal is absent. This is in effect a more conservative approach which aims to prevent the exclusion of signals to which an experiment is not sufficiently sensitive (*i.e.* a false exclusion, which can be caused by downward statistical fluctuations in the total background). By definition CL_s is always greater than CL_{s+b}, meaning that a test based on the former will yield a higher limit. A CL_{s+b} value of 0.05 is equivalent to an exclusion with a confidence level of 95 %.

For the purposes of Higgs physics, the relationship between the discovery/exclusion thresholds (quantified in terms of confidence levels) and the Higgs mass is typically shown on plots such as Figure 6.1. The plot shows the expected Higgs cross-section which can be excluded at a confidence level of 95 %, expressed as a multiple of the standard model cross section, together with the 1σ and 2σ errors on the expected confidence level. The observed CL_s, as calculated from data, is then compared with the expectation. On such a plot, a discovery at a certain mass will be visualised as a large "bump" of the observed data (*i.e.* an excess of events) with respect to the background-only expectation.

6.3 Basic Analysis Methods

6.3.1 Cuts-Based Analyses

The oldest and simplest method used by physicists in selecting interesting events is simply to impose cuts in order to define the relevant phase-space regions. Most simply, one chooses cuts such that relevant observables will fall within given ranges,

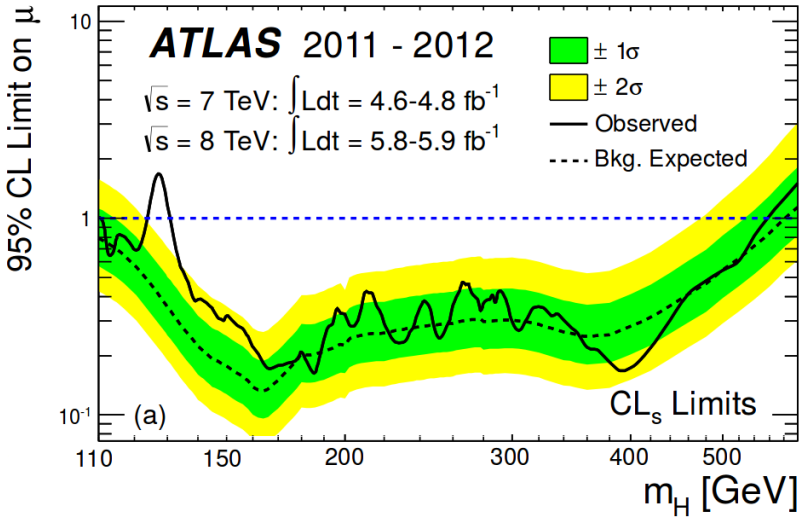


Figure 6.1: A plot showing the Higgs discovery by the ATLAS collaboration using the LHC Run 1 dataset. At a m_H of approximately 125 GeV, the combined Higgs signal strength μ clearly exceeds the background-only expectation. Adapted from Ref. [6].

in a manner not dependent on other observables (so-called *rectangular* cuts); this can be difficult, and requires good knowledge of the relationships between the different variables in relevant events. One can also construct a discriminating variable from a linear combination of observables, in accordance with Fisher's method [107].

The basic procedure of a cuts-based analysis starts with requiring events to have the relevant *topology*. This means that there must be certain numbers of observable final-state objects (photons, electrons, jets, etc.) which meet certain requirements, such as exceeding a certain p_{T} threshold, or falling within a certain pseudorapidity region. Many of these cuts are imposed at the level of object identification and reconstruction (since, for example, extremely soft particles are unlikely to be of interest). Subsequently, kinematic cuts are imposed in order to reject background events as required; these can take the form of requiring a minimal missing transverse energy (denoted $E_{\mathrm{T}}^{\mathrm{miss}}$ or MET) or a maximal angular separation (ΔR) between certain pairs of particles.

Whereas the principles underlying cuts-based analyses are straightforward, the technique has a number of important weaknesses. Firstly, the procedure is inexact, and somewhat arbitrary, since the very large number of possible combinations of different cuts implies that it is extremely difficult to ascertain whether a given set of cuts will produce an optimal selection of events for analysis. This difficulty typically increases

with the complexity of the event topology under consideration. The marginal effect of each additional cut should be carefully studied in isolation in order to understand its consequences for the analysis; the cumulative effect of the cuts as they are applied gives rise to the "cut flow" of the analysis. Moreover, care must be taken to ensure that the cuts are chosen without bias: by looking at data, it is possible that one could simply choose cuts so as to produce a desired result, potentially causing the appearance of signal where none exists; this can especially problematic when the cuts are such that few events are ultimately selected, owing to the proportionately larger effect of statistical anomalies in small samples. Blinding is typically employed to mitigate the effects of bias: physicists will deliberately avoid examining certain regions of data while they are choosing which cuts to apply, so as to minimise the risk of choosing cuts that are overly-specific to the particularities of a single data set, or likely to give a biased result.

6.4 Multivariate Neural Network Analyses

6.4.1 Theory of Neural Networks

When considering problems in multivariate data analysis, it is often found that cut-based analysis methods are less than optimal. With many variables to consider in order to define phase-space regions, it can be difficult to determine a priori what the most relevant variables to a given analysis will be, and hence the ability of physicists to maximise their sensitivity to signals is hampered. Thus it is desirable to implement analysis methods which do not depend on detailed knowledge of how the sensitivity relates, for example, to particular choices of cuts on observables, which are in any case somewhat arbitrary.

There are many different multivariate computational methods which find use in particle physics, such as boosted decision trees [108, 109], automated likelihood estimation algorithms [110], support vector machines [111], and artificial neural networks [112, 113]. The great strength of such learning-based approaches is that they are able to model the behaviour of functions (*i.e.* the mapping between the input data set and the desired outputs) which might be unknown in advance and extremely difficult to describe as a set of formal rules [114].

With the rapid increase in available computational resources, these learning-based methods have seen growing use. The method that will be used in the analysis work described in this thesis is that of the artificial neural network (ANN), more

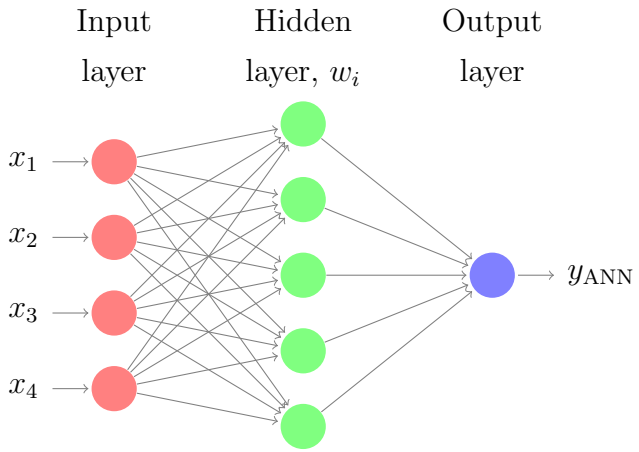


Figure 6.2: A diagram illustrating the structure of a simple feedforward artificial neural network, featuring one hidden layer and one output node. The inputs x_i are assigned weights w_i; the weighted sum then gives the output y_{ANN} in accordance with Equation (6.16).

specifically a multi-layer perceptron (MLP) network [115]. An ANN is a method of implementing some of the functionality of an animal's nervous system (*i.e.* the ability of the brain to learn information, in a very general sense) within a software framework. ANNs can be used either to describe decision boundaries (termed *classification*) or to model the underlying dynamics of the data set, with a view to understanding the functional relationship between the inputs and a desired target value (for instance, an experimental sensitivity – this approach is referred to as *regression*). An important mathematical result in the theory of neural networks is the Cybenko theorem [116], which proves the general ability of MLPs (specifically defined as feed-forward networks with a single finite hidden layer) to approximate the behaviour of an arbitrary continuous real-valued function, where a suitable activation function (a sigmoid function such as a logistic function or an hyperbolic tangent) is chosen for the network. This implies that MLPs have an extremely wide range of applications, as in principle they can be used to describe the behaviour of any multivariate function, including in cases where such functional behaviour cannot be described analytically.

6.4.2 Structure of a Multi-Layer Perceptron

The structure of a neural network consists of layers of highly-interconnected processing elements termed *neurons* or *nodes*: an input layer, at least one hidden layer where the processing occurs, and an output layer – this structure is illustrated in

Figure 6.2. On input, each variable is assigned to a neuron. The process of training then adjusts the weights assigned to the connections between these neurons by running a sequence of training cycles, in each of which the weights are slightly altered, with the goal of optimising the overall sensitivity of the network to patterns in the data. Each intermediate neuron will activate if the weighted sum of its inputs exceeds a certain threshold, and its activation is used in turn as an input into the next layer of processing, until output is reached. The weighted output y_{ANN} is given by the formula

$$y_{\mathrm{ANN}} = \sum_{j=1}^{n_{\mathrm{h}}} y_j^{(2)} w_{j1}^{(2)} = \sum_{j=1}^{n_{\mathrm{h}}} \tanh\left(\sum_{i=1}^{n_{\mathrm{var}}} x_i w_{ij}^{(1)}\right) \cdot w_{j1}^{(2)} \,, \tag{6.16}$$

where the activation function is taken to be a hyperbolic tangent, for inputs x_i, weights w_{ij}[(ii)], and where n_{h} denotes the number of hidden layers and n_{var} denotes the number of input variables [107]. The MLP is a specific case of a feedforward network, meaning that each layer of neurons is connected directly to the neurons in the next layer; this category excludes network topologies with directed cycles of neurons (*i.e.* recurrent networks).

6.4.3 Neural Network Training

The mathematical principle underlying the training is that of spectral decomposition; the goal is to find the principal axes of the input data set, considered as a structure in an n-dimensional configuration space. By reweighting the connections between its neurons, the neural network is able to change the orientation of this structure.

The ANN is an implementation of supervised learning, meaning that it aims to improve its performance at computational tasks by gaining experience of relevant data, and its learning is supported by the availability of "correct" responses to certain of its inputs – the network is given feedback on its responses during training. The objective is for the network to be able to generalise from this trained data in order to give similarly correct responses in situations where the outcomes are not known in advance; without such "supervision", a learning algorithm could at most observe patterns, without being able to produce appropriate responses to these patterns [117].

[(ii)] Superscript numeral (1) denotes the weights used in the hidden-layer sum over the inputs, while (2) denotes the weights used to sum over the hidden nodes to give y_{ANN}.

The method that is employed is that of backpropagation, meaning that the output of some error function is propagated "backwards" to the neural network, constituting a kind of feedback system. The "wrongness" of the neural network's modelling of its inputs is quantified by means of a loss function (or "loss matrix" – effectively a kind of negative utility function). The objective of training can then be formally defined as a constrained minimisation of the loss function, using the method of gradient descent, *i.e.* by computing the gradient of the loss function and then varying its inputs in the direction of the negative gradient, towards the minimum [118, p. 27]. A typical error function has the form

$$E_a = \sum_{a=1}^{N} \frac{1}{2} \left(y_{\mathrm{ANN},a} - \hat{y}_a \right)^2 , \tag{6.17}$$

where y_{ANN} is the weighted ANN output and $\hat{y}$ is the "correct" response (typically this would be set to -1 for background events and +1 for signal events) [107]. And thus, by the method of gradient descent, on each iteration the weights are modified according to the rule

$$w^{\rho+1} = w^{\rho} - \eta \nabla_w E, \tag{6.18}$$

where the parameter η denotes the learning rate (*i.e.* the size of the step taken per iteration of training) and ∇_w denotes the gradient with respect to the weights w; the index ρ represents the number of training cycles [107].

While a naive approach might lead one to think that a larger amount of training will invariably produce a more useful network, it is in reality very important to choose an optimal number of training cycles, as training can be thought of as a tradeoff between the neural network's "understanding" of its inputs and its generality – this is also referred to as the bias-variance tradeoff [119]. An insufficiently general (or overtrained) neural network will give too high a weighting to the specific properties of the data set on which it was trained, and it will thus become less capable of identifying similarities in other data sets. This is especially a problem in complex analyses where the number of necessary free parameters is smaller than the number of parameters afforded by the network, allowing the network to "over-fit" to its data set (for example, networks which have been overtrained on several slightly different data sets could give inconsistent output). These problems can be circumvented by cross-validating the trained network against a control sample to test for divergence in the weights; most commonly, this is done by dividing the input data set into

separate subsets for training and validation [118].

6.4.4 Data Analysis Software

ROOT

An important data analysis tool in many high-energy physics analyses, ROOT is an object-oriented framework which can be thought of as a library of inter-related classes of C++ objects intended for use in data analysis and visualisation. Basic and widely used functions of ROOT include the creation of objects such as histograms, which may be plotted and presented in a wide variety of ways (*e.g.* contour plots and surface plots) so as to aid physicists in their understanding of their data. ROOT may be run interactively with the aid of a command-line C++ interpreter such as CINT or Clang, and its objects and functions may be imported into python by means of PyROOT. Equally, it is possible to include ROOT libraries in compiled C++ code.

TMVA

The neural network used in the analysis presented in this thesis was implemented in the Toolkit for Multivariate Data Analysis (TMVA) [114], a software package which integrates with the general-purpose analysis framework ROOT. This package provides a wide range of multivariate algorithms, covering functionality such as likelihood estimation, boosted decision trees, cut optimisation, as well as three different implementations of a multilayer perceptron network. It accepts input data in the form of ROOT trees, over which training can be run.

Running as a classifier, TMVA's multilayer perceptron attempts to learn how to disentangle signal events from a specified set of background events, as defined during the training procedure and using the available event data. Typically it will assign higher classification values (often referred to as *neural net scores*) to events which it considers to resemble the signal more closely, whereas events resembling those from background processes are assigned lower classification values. In terms of the description of cuts outlined in Section 6.3.1, one can think of TMVA as being asked to discover an optimal decision boundary, while considering that such a boundary might not be readily described via a simple parametrisation (such as rectangular cuts or a linear function of observables). In principle, the MLP has complete freedom to explore the space of possible decision boundaries, in order to select the boundary

which offers optimal discriminating power (which can be thought of as the best ratio of signal acceptance to background rejection).

NeuroBayes

Another software package which may be used to implement MLPs is NeuroBayes, a proprietary suite of tools for carrying out advanced multivariate analyses. As its name implies, it combines the strengths of neural network analysis techniques with Bayesian statistics, and it is intended to be overtraining-safe (by contrast, TMVA typically requires a certain amount of oversight from the end-user to ensure that overtraining does not occur) [120].

NeuroBayes implements training (typically run over simulated data) inside an environment called the NeuroBayes Teacher, producing a collection of weighted nodes which may be passed downstream to the following phase of the analysis, which is implemented inside the NeuroBayes Expert system. This system combines the results of training with real data to carry out a Bayesian hypothesis testing (*e.g.* testing the real data against signal-plus-background and background-only hypotheses).

RooFit

RooFit is a library of analysis functions for inclusion with ROOT, designed broadly to model the expected distributions of events and event data [121]. These models may be used to carry out fits, such as unbinned maximum likelihood fits, from which exclusions can be calculated.

In such an event data distribution, each event is considered to be a discrete occurrence, with some number of associated variables; this is characteristic of Poisson statistics. In the most general sense, then, the task of RooFit is to model such distributions in terms of probability density functions of the form $F(x, p)$ where x represents an observable and p represents a function parameter.

As with ROOT itself, RooFit is conceptually object-oriented, meaning that each distinct concept in an analysis (such as a variable, an event, a function or a distribution) is represented as an object in C++; these objects are characterised not by the functions they perform in the context of a given analysis but rather by the data type which they represent; RooFit aims to create a one-to-one mapping of mathematical and statistical concepts such as variables and functions onto its classes.

A key concept in performing a fit with RooFit is that of the *workspace* – this is a

file containing all the likelihood functions for a particular channel. These functions are combined to create a combined probability model, including the effects of nuisance parameters corresponding to systematic effects. In this way it is possible to combine results from multiple channels, with appropriate coordination to allow for the differing effects of systematics (*e.g.* detector effects) on the different channels. Specifically, it is possible to use the HistFactory tool to create a workspace based on template histograms, which serve as inputs to a fit.

Limit-setting code

The final limit setting and exclusion is carried out using code [122] which takes a RooFit workspace as input and performs a constrained minimisation using Minuit [123]. This code works by implementing the iterative formula

$$\mu_{N+1} = \frac{\mu_N - (p_\mu - (1 - \mathrm{CL}))}{\frac{\mathrm{d}p_\mu}{\mathrm{d}\mu}} \tag{6.19}$$

to obtain an excluded signal strength μ_N for some number of iterations N, where p_μ is the p-value corresponding to the signal strength μ and CL is the confidence level. This method is a proxy to running a vast number of pseudoexperiments, following the asymptotic limit-setting procedures outlined in Ref. [104]. In the code, Equation (6.19) is approximated as

$$\mu_{N+1} \approx \frac{\mu_N - (p_\mu - (1 - \mathrm{CL}))}{\frac{p_{\mu(1+\epsilon)} - p_\mu}{\mu(1+\epsilon) - \mu}}\,, \tag{6.20}$$

where the parameter ϵ defaults to 0.05. The code computes limits on the strength of a positive signal (equivalently, the product of cross section and branching ratio, $\sigma \times \mathcal{B}$) for a given value of m_H and a given Higgs decay mode: the observed limit (*i.e.* using real data), the median expected limit (derived from simulated or "Asimov" data) and the 1- and 2-sigma limits in both directions.

6.5 Summary

Several of the statistical methods and analysis techniques employed in high-energy physics have been introduced and described in outline, with a focus on the analysis framework used for the work presented in this thesis, namely an artificial neural network implemented in TMVA. The principles underlying the concepts of sensitivity,

discovery and exclusion – key to high-energy physics analyses – have been explained, with a particular view towards their applications to Higgs physics.

Chapter 7

Search for $t\bar{t}H, H \to b\bar{b}$

"Lasciate ogni speranza, voi ch'entrate."

— Dante, *Divina Commedia*

7.1 Introduction

The analysis presented here was performed using ATLAS pp collision data from 2012, with a centre-of-momentum energy of 8 TeV. Specifically, it is a search for Higgs production in association with the top quark, with the Higgs decaying to a bottom pair, and the two top decays producing respectively a pair of jets and a single isolated lepton (*i.e.* the semileptonic channel, as described in Chapter 2). For the purposes of this analysis, the lepton is allowed to be either an electron or a muon, including leptons arising from leptonic tau decays. The regions of events under consideration are defined with reference to the total number of jets and the number of b-tagged jets they contain. Before the analysis event selection is implemented, there are 439 725 data events with an electron and 525 903 data events with a muon, giving a total of 965 628 events.

7.2 Object Definitions

In this search for semileptonic $t\bar{t}H$ production, the most important physics objects in each event are a hard lepton (which may be either an electron or a muon), jets, and b-jets. Each of these categories of object is required to satisfy certain criteria for the purposes of this analysis. The object definitions and selection criteria used

in this analysis broadly follow those of the ATLAS $t\bar{t}H, H \to b\bar{b}$ working group as outlined in Ref. [93]; they were chosen, following studies undertaken by the group, so as to optimise the sensitivity of the analysis by maximising the yield of signal events in those analysis regions which are designated as signal-enriched.

Electrons are identified by associating clusters of energy deposited in the electromagnetic calorimeter to inner detector tracks; each such cluster is required to have $|\eta_{\text{cluster}}| < 2.47$, excluding the calorimeter transition region $1.37 < |\eta_{\text{cluster}}| < 1.52$ [124]. Since it is possible for electrons to be produced outside of the hard-scattering process (*e.g. non-prompt* electrons originating in hadron decays), an isolation cut is imposed on electron candidates. This cut is based on a cone of $\Delta R = 0.2$ around each candidate; the total transverse momentum contained in the cells falling within this cone (excluding those associated directly to the candidate itself, and correcting for energy leakage) is then used as a cut variable, yielding a cut of 90 % efficiency. A further isolation requirement is imposed on the sum of track transverse momenta within a cone of $\Delta R = 0.3$ around the electron candidate ($p_{\text{T}}^{\text{cone30}}$), also yielding an efficiency of 90 %. The final requirement is that the track's longitudinal impact parameter z_0, as measured with respect to the primary vertex of the event, must be less than 2 mm.

Muon identification is performed by associating track segments in the muon spectrometer to inner detector tracks; muon candidates must satisfy a pseudorapidity cut of $|\eta| < 2.5$. There is an additional isolation requirement of $\Delta R > 0.4$ from any identified jet. There is also a requirement that the scalar sum of track transverse momenta (excluding those of the muon tracks) inside a cone with p_{T}-dependent radius $\Delta R = 10\,\text{GeV}/p_{\text{T}}^{\mu}$ must be less than 5 % of the transverse momentum of the muon. Tracks in the inner detector are required to be well-reconstructed, and the requirement on track longitudinal impact parameter being less than 2 mm with respect to the primary vertex is also imposed upon muon candidates.

The reconstruction of jets is performed by associating clusters of energy inside the calorimeters to one another [125]; specifically, the anti-k_t algorithm is used with a cone radius parameter of $R = 0.4$ [86]. Before jet-finding is performed, a local cluster calibration scheme is used to correct for effects such as dead calorimeter material and energy leakage from calorimeter regions [126]. Calibration is performed using factors dependent on energy and pseudorapidity, which are derived from simulation and corrected using *in situ* techniques [127]. Finally, jet candidates must satisfy the requirements $p_{\text{T}} > 25\,\text{GeV}$ and $|\eta| < 2.5$. For jets with $p_{\text{T}} < 50\,\text{GeV}$ and $|\eta| < 2.4$, there is an additional requirement that the scalar sum of transverse momenta of

the tracks originating at the primary vertex and matched to the jet must be at least 50 % of the total p_{T} of all tracks matched to the jet. This requirement aims to reduce contamination due to pileup [62]. Since there is no distinction at the calorimeter level between identified electrons and energy deposits associated to jets, one must compensate for potential double-counting of electrons as jets by discarding the closest jet in the case where there are jets within $\Delta R < 0.2$ of an identified electron.

b-tagging is implemented using a multivariate algorithm which takes as its inputs the impact parameters of the displaced tracks and the topological properties of the reconstructed secondary and tertiary vertices within the jet [128]. The b-tagging working point used for this analysis yields a 70 % b-tag efficiency and a 1 % mistag rate, as determined using b-tagged jets satisfying p_{T} and $|\eta| < 2.5$ in simulated $t\bar{t}$ events [93].

7.3 Event Selection

This analysis was performed using proton-proton collision data with centre-of-momentum energy $\sqrt{s} = 8\,\mathrm{TeV}$, collected by the ATLAS experiment in 2012; all events used in the analysis are required to contain a single hard lepton (electron or muon) trigger under stable beam conditions, with all detector subsystems operational – this gives a total integrated luminosity of $20.3\,\mathrm{fb}^{-1}$.

For electrons and muons, two sets of trigger p_{T} thresholds – 24 GeV and 60 GeV for electrons and 24 GeV and 36 GeV for muons – are used to maximise the total efficiency; these are combined using a logical OR. This approach is adopted because the triggers at low p_{T} have lepton isolation requirements which lower efficiency at higher p_{T}. Notably, the trigger selection criteria are less strict than the object reconstruction criteria which are imposed downstream. The muon trigger efficiency is approximately 0.7 [129].

Events must have exactly one identified hard lepton, which may be an electron or a muon. This lepton is required to have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, and it must match the lepton reconstructed by the high-level trigger to an accuracy of $\Delta R < 0.15$. Notably, events with multiple identified leptons which also match the selection criteria for the $t\bar{t}H$ dilepton channel have been removed in order to minimise statistical overlap between the lepton-plus-jets and dilepton channels. The additional requirement $E_{\mathrm{T}}^{\mathrm{cluster}} > 25\,\mathrm{GeV}$ is imposed on electrons.

The trigger imposes a further requirement that events have at least one reconstructed vertex with at least five associated tracks consistent with the beam collision region in the (x, y) plane. If there are several such vertices in an event, the vertex with the highest square-transverse-momentum sum over tracks is taken to be the hard scattering primary vertex.

Moreover, the event must also contain at least four reconstructed jets which satisfy the appropriate quality and kinematic criteria for this analysis (as described in section 7.2). At least two of these jets must be b-tagged.

$E_{\mathrm{T}}^{\mathrm{miss}}$ is defined and computed in accordance with the dynamical likelihood methods outlined in Ref. [130]. An additional requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 25\,\mathrm{GeV}$ is imposed with the intention of reducing the multijet background, as described in Appendix B.

7.4 Analysis Regions

Following event selection, events are categorised according to a scheme of nine non-overlapping regions, defined by the event's jet multiplicity (where only reconstructed jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, including b-tagged jets, are counted) and the number of b-tagged jets in the event – these nine regions are represented in Table 7.1. As shorthand, the region corresponding to x jets, of which y are b-tagged jets, will be referred to using the notation (xj,yb), contracted to (x, y). Note that "6j" and "4b" are always *inclusive* (*i.e.* denoting "6 (4) or more (b)-jets") while other values are *exclusive* (*e.g.* "5j" denotes "exactly 5 jets"). In this analysis, each of these nine regions is analysed separately; the results of these analyses are then statistically combined to produce a maximally sensitive result. Firstly, three *signal-enriched* regions are identified: these are (5,4), (6,3) and (6,4). These are the regions in which the signal is most prominent, and thus the majority of the sensitivity of the analysis is expected to derive from them. For this reason, the neural network is trained separately in each of these regions in order to maximise the signal-background separation (and thus the sensitivity). These regions were also blinded during the initial stages of the analysis, while the signal-depleted regions were designated as control regions – these were the regions (4,2), (4,3), (4,4), (5,2), (5,3), and (6,2). By including these control regions – which are dominated by different background processes – in the fit, additional information is gained about the nature of the systematic uncertainties.

	n_j		
n_b	4	5	≥ 6
2	BR	BR	BR
3	BR	BR	SR
≥ 4	BR	SR	SR

Table 7.1: Table outlining the analysis regions used in the Standard Model analysis. "BR" denotes background (signal-depleted) regions while "SR" denotes signal-enriched regions.

7.5 Analysis Software Framework and Workflow

The analysis described in this thesis was performed inside an analysis framework called GlaNtp, designed with the intention of streamlining all the stages in the analysis, including reading and reweighting the input distributions, training an instance of TMVA, producing stacked plots and computing a final exclusion. The analysis workflow is summarised in Table 7.1.

Firstly, the ATLAS-format input files (signal and background Monte Carlo, and data) need to be prepared using a stripper and ntuple-modification tool which removes some superfluous branches from each file, as well as adding some additional branches containing cutword information necessary for the event selection which is implemented in the following stages.

Within those regions in which a neural network training is performed, the analysis proper then begins by running software to reweight the events so that they are represented in the appropriate proportion within the sample, and to use the information added by the stripper tool to classify the events according to a predefined scheme of analysis cuts, such that only those events within the desired region (as specified in the appropriate analysis software steering files) survive this cut. These reweighted Monte Carlo distributions, signal and background, are combined into one input ROOT file per analysis region.

These ROOT files are then passed to a second program which trains an instance of the TMVA in accordance with steering files specifying the signal and background processes to train over, as well as the training variable set. The internal structure of the neural network (number of layers and number of nodes per layer, typically

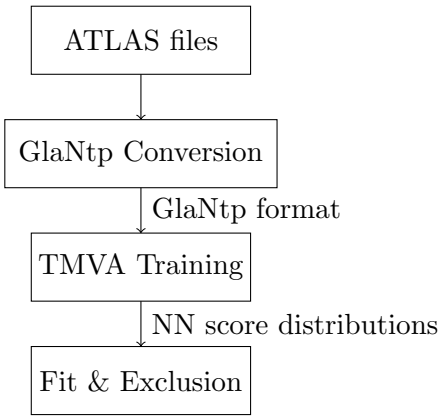


Figure 7.1: Diagram summarising the workflow of the GlaNtp analysis. The output format of each stage is indicated by the label on the arrow leading to the following stage.

expressed in terms of the number of inputs) and the number of training cycles may also be specified to attain an optimal result. The neural net structure used for this analysis has two hidden layers of $N + 1$ and N nodes respectively (where N is the number of input variables); 500 training cycles are used, as this number was found to produce convergence without risking overtraining.

The outputs of these two stages are then passed to the a third program to create ROOT template files combining the pre- and post-training information; one template is created for each process. These templates serve as the building blocks for the stacked plots of the input variable distributions and TMVA score, as well as inputs to the fit process. GlaNtp creates, for each of the nominal values and systematic variations in each analysis region, a single ROOT file containing the appropriately binned fit template histograms for each process, with the signal and backgrounds combined and weighted according to their cross sections (as detailed in Section 7.6) – these template files are used as an input to a RooFit workspace, which is in turn used as input to HistFitter [131] and the asymptotic limit-setting code as described in Section 6.4.4 to calculate the final exclusion.

7.6 Background and Signal Modelling

The main contribution to the background in the standard model $t\bar{t}H$ analysis is $t\bar{t}$+jets production (where the additional jets include both light and heavy flavours), with further important contributions coming from single-top production, production of a vector boson ($W^{\pm}, Z$) in association with jets, diboson production (WW, ZZ, WZ),

and production of a vector boson in association with a $t\bar{t}$ pair (denoted $t\bar{t}+V$). Moreover, there are multijet contributions due to misidentification of jets or photons as electrons, and due to non-prompt leptons arising from the decays of $b-$ and $c-$ hadrons.

$W+$jets and $Z+$jets events, together with diboson production events, are simulated in Alpgen 2.14 [132] using the CTEQ6L1 PDF set [89]; fragmentation and parton shower evolution are simulated using Pythia 6.425 [133] in the case of $W+$jets, and Herwig 6.520 [134] for $Z+$jets and dibosons. The MLM matching technique is used to match jets and partons, in order to avoid double-counting between the matrix-element calculation and the fragmentation; this technique employs the ALPHA algorithm [135] to perform matrix element calculations to extract colour information for multiparton processes [136]. The $W+$jets and $Z+$jets backgrounds are estimated via p_{T}-reweighting of the Monte Carlo to account for differences in the p_{T} spectra of data and Monte Carlo; the heavy-flavour fractions of these background processes are adjusted to match up with the relative frequencies in real data of events containing a Z and 0 or 1 b-tag [137].

The $W+$jets samples are generated with up to five additional partons (separately for each of $W+$light jets, $Wb\bar{b}+$jets, $Wc\bar{c}+$jets and $Wc+$jets, where "light jets" denotes jets not identified as originating from a b or c quark). The overlap between $WQ\bar{Q}$ events (where $Q = b, c$ represents heavy quark flavours) arising from the matrix element calculation, and those arising from the showering of events containing light jets, is removed by considering the ΔR between the heavy-flavour quarks: if $\Delta R(Q, \bar{Q}) > 0.4$ then the matrix-element prediction is used; otherwise the prediction from the parton shower evolution is used. $Z+$jets events are generated according to the same procedure, and both $W+$jets and $Z+$jets events are normalised to their respective theoretical NNLO inclusive cross-sections [138]. The diboson events are generated with up to three additional partons; they are normalised to their theoretical NLO cross-sections [139].

The $t\bar{t}+$jets background samples are generated with Powheg [92, 140, 141], using the CT10 PDF set with a nominal top quark mass of $m_t = 172.5\,\mathrm{GeV}$, which is interfaced to Pythia 6.425, using the CTEQ6L1 PDF set and the Perugia2011 underlying-event tune. The sample is normalised via an NNLO-QCD calculation, which includes the resummation of NNLL soft gluon terms using top++ 2.0 [142–148], giving a total cross-section of $253^{+15}_{-16}\,\mathrm{pb}$ at $\sqrt{s} = 8\,\mathrm{TeV}$, where the systematic uncertaintly includes errors arising from the PDF and α_s, as well as the mass of the top quark. The uncertainties in the PDF and α_s were calculated according to the

PDF4LHC prescription [149], using the MSTW2008 68 % NNLO [150, 151], CT10 NNLO [152, 153] and NNPDF2.3 5f FFN [154] PDF sets, added in quadrature to the scale uncertainty. The sample is generated inclusively, but events are characterised by the flavour content of any jets they contain which do not originate in the decay of the $t\bar{t}$ system. Jets are reconstructed using the anti-k_t algorithm with radius parameter $R = 0.4$ and the requirements that $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$. Those events with at least one such jet matched to a b-quark not originating from the decay of the $t\bar{t}$ system are labelled $t\bar{t}b\bar{b}$ events, and those events which are not labelled $t\bar{t}b\bar{b}$ and have at least one jet matched to a c-quark not originating from the decay of the $t\bar{t}$ system (via a $W^{\pm}$ boson) are labelled $t\bar{t}c\bar{c}$ events. These $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ events are collectively denoted $t\bar{t}+\mathrm{HF}$, where "HF" represents "heavy flavours". All those $t\bar{t}+$ jets events which are not classified as $t\bar{t}+\mathrm{HF}$ are collectively referred to as $t\bar{t}+$ lightjets.

In order to optimise the agreement between simulation and data for the $t\bar{t}$ sample, a correction is applied to all simulated $t\bar{t}$ events which is based on the ratio of differential cross sections at $\sqrt{s} = 7\,\mathrm{TeV}$ of the data and simulation; this correction is a function of the top quark p_{T} and the p_{T} of the $t\bar{t}$ system. This correction procedure improves the agreement between simulation and data in terms of the total jet multiplicity and the jet p_{T}.

The single-top samples are generated in POWHEG with the CT10 PDF set; these correspond to the Wt and s-channel production mechanisms. For the t-channel, the AcerMC v3.8 LO generator is used with the MRST LO** PDF set. These samples are interfaced to PYTHIA 6.425 with the CTEQ6L1 PDF set and the Perugia2011C underlying-event tune. Overlaps between Wt and $t\bar{t}$ are removed, and the single-top samples are normalised to their theoretical NNLO cross-sections using the MSTW2008 NNLO PDF set [155].

The $t\bar{t}V$ samples are generated in MADGRAPH 5 using the CTEQ6L1 PDF set, interfaced with PYTHIA 6.425 using the AUET2B tune for parton shower evolution. These samples are normalised to their theoretical NLO cross-sections.

The $t\bar{t}H$ signal is modelled inclusively with respect to Higgs decays, using the HELAC-Oneloop package [156] to give NLO matrix elements; these are then showered using POWHEG BOX [142] as an interface to PYTHIA 8.1, with the PDF set CTEQ6L1 and the AU2 underlying event tune. The resulting Monte Carlo samples are referred to as POWHEL samples. The PDF set used is CT10NLO, with factorisation and renormalisation scales $\mu_0 = \mu_{\mathrm{F}} = \mu_{\mathrm{R}} = m_t + m_H/2$. For the purposes of this analysis, $m_H = 125\,\mathrm{GeV}$ is taken to be the Higgs mass, $m_t = 172.5\,\mathrm{GeV}$ is taken

to be the top quark mass, and the the total $t\bar{t}H$ cross section and Higgs branching ratios are taken from theoretical NLO and NNLO calculations [37].

All simulations use PHOTOS 2.15 [157] to simulate photon radiation and TAUOLA [158] to simulate τ decays. In every simulation involving HERWIG, JIMMY 4.31 [159] is used to simulate the underlying event. Events from minimum bias interactions are simulated in PYTHIA 8.1 using the MSTW2008 LO PDF set and the A2 tune [90] [91]; these events are overlaid on the simulated signal and background events according to the luminosity profile of the recorded data. These pile-up interactions are simulated within the same buch crossing as the hard-scattering process, and within neighbouring bunch crossings.

All simulated samples are passed through a simulation of the geometry and response of the ATLAS detector. This simulation either uses GEANT4 [160, 161] consistently throughout, or else makes use of a fast simulation of the calorimeter response. The samples are then processed using the same reconstruction software as the data. Simulated events are corrected so that the object identification efficiencies, energy scales and energy resolutions match those of control samples made from data.

In those analysis regions requiring high jet multiplicity or high b-tag multiplicity, large fluctuations can arise in the simulated distributions due to low statistics. This leads in turn to a higher statistical uncertainty, and it can lead to erroneous calculations of the shape systematic uncertainties, with a corresponding reduction in sensitivity of the analysis and the potential introduction of a bias, depending on how these fluctuations affect the simulation relative to the data in signal-enriched analysis regions. This problem is reduced by adopting a modified approach to b-tagging: rather than simply applying the b-tagging algorithm to tag jets, a tag-rate function (TRF) is constructed, parameterised in terms of jet p_{T} and η, and this function is used to weight all events in the sample by the probability that they contain a given number of b-jets. With this function, all events in the simulated sample may be used to predict the normalisation and shape post-tagging.

7.6.1 The Tag Rate Function

In analysis regions which require high jet multiplicity or a high number of b-tagged jets, the number of available events in the Monte Carlo samples is greatly reduced, leading to low statistics in these regions. This means that there are likely to be large statistical fluctuations on bin content in such regions, with a corresponding reduction in the sensitivity of the analysis.

The effect of this problem is reduced by applying the tag rate function (TRF) to all Monte Carlo samples, in order to allow the events before b-tagging to be used to predict the normalisations and shapes of distributions after b-tagging.

7.7 Systematic Uncertainties

A typical high-energy physics analysis is subject to a wide variety of sources of systematic uncertainty, arising from detector effects, event reconstruction, and signal and background Monte Carlo modelling. These systematic uncertainties are modelled in Monte Carlo as affecting the shape and the normalisation of the simulated signal and background distributions, causing the predicted content of each bin in the analysis to fluctuate up or down with respect to the data. The net effect of these fluctuations on the fit and limit setting can be to exaggerate or diminish the appearance of signal in the data. For the purposes of this analysis, the various sources of systematic uncertainty are assumed to be statistically independent (*i.e.* uncorrelated); for each source of uncertainty, the correlations are taken to be identical for each process and channel. The estimated luminosity has a relative uncertainty of 2.8 % [162], which is applied as an overall uncertainty to all processes (signal and background) that are simulated in Monte Carlo.

The various types of systematic uncertainty considered in this analysis are shown in Table 7.2, which indicates whether each type is considered as an uncertainty on the shape of the discriminant spectra, or on the normalisation, or both; together with their respective numbers of constituent components (*i.e.* the number of independent sources of uncertainty that are considered for the purposes of this analysis to contribute to each kind of systematic uncertainty).

7.7.1 Leptons

In this analysis, five independent sources of systematic uncertainty are considered for electrons, and six for muons. The lepton selection contributes to the overall systematic uncertainty via the trigger, identification, isolation, reconstruction, lepton momentum scale, and resolution.

Lepton momentum scale and resolution are validated by comparing distributions, in simulation and data, of reconstructed $Z \to l^+l^-(l = e, \mu)$ and $J/\psi \to l^+l^-$ invariant masses. E/p studies in $W \to e\nu$ events provide an additional means of verifying

Source of Systematic Uncertainty	Type	No. Components
Luminosity	N	1
Electrons	SN	5
Muons	SN	6
JES	SN	22
JVF	SN	1
JER	SN	1
Jet reconstruction efficiency	SN	1
b-tagging efficiency	SN	6
c-tagging efficiency	SN	6
Light-jet tagging efficiency	SN	12
$t\bar{t}$ cross-section	N	1
$t\bar{t}$ modelling – p_{T}-reweighting	SN	9
$t\bar{t}$ modelling – parton showering	SN	2
$t\bar{t}$+HF – normalisation	N	2
$t\bar{t}$+HF – HF reweighting	SN	2
$t\bar{t}$+HF – generator	SN	5
W+jets normalisation	N	3
W p_{T}-reweighting	SN	1
Z+jets normalisation	N	2
Z p_{T}-reweighting	SN	1
Multijet normalisation	N	3
Multijet shape dilepton	S	1
Single top cross-section	N	1
Dibosons cross-section	N	1
$t\bar{t}V$ cross-section	N	1
$t\bar{t}H$ modelling	SN	2

Table 7.2: Table of the sources of systematic uncertainty which are considered in this analysis. For each source of uncertainty, the number of components is given, along with whether that uncertainty is considered to affect the normalisation ("N") of the fit templates, or both the shape and the normalisation ("SN") .

the accuracy of electron momentum scale and resolution. Corrections are applied to compensate for small data-simulation discrepancies in lepton energy scale and resolution; for electrons, these corrections are applied to both data and simulation in the case of the energy scale, and for the case of energy resolution they are applied solely to the simulation. For muons both corrections apply only to simulation, and moreover there are additional sources of uncertainty arising from the muon spectrometer and the tracking systems.

The efficiencies of lepton reconstruction and identification show small discrepancies between data and simulation – to correct for this, scale factors are applied to the simulation. These scale factors are computed in accordance with the tag-and-probe method, comparing data and simulation for $Z \to l^+l^-$. The total uncertainty on lepton identification, trigger, isolation and reconstruction introduces an uncertainty of 1.3 % into the event yields.

7.7.2 Jet Energy Scale (JES)

The total JES uncertainty is represented as arising from 22 independent sources, which are taken to be uncorrelated. These different sources have varying functional dependence on the p_{T} and η of jets. The estimated JES uncertainties derive from test-beam data, collision data, and simulation [127, 163–167]. These 22 components are listed in full in Table D.1 on page 147.

7.7.3 Jet Vertex Fraction (JVF)

The jet vertex fraction (JVF) is defined as the fraction of track p_{T} associated to a jet which originates in a hard-scattering process [129]. Data and simulation are compared in $Z \to (l^+l^-) + j$ events to measure the value of the per-jet efficiency which is necessary to satisfy the JVF requirement. Events enriched in jets originating from hard scattering are selected separately from events enriched in jets originating from separate proton interactions within the same bunch-crossing. The uncertainty corresponding to this efficiency is calculated by observing the effect of varying the JVF cut value up and down by 0.1.

7.7.4 Jet Energy Resolution (JER)

The JER is measured separately for data and simulation, and then compared to estimate the systematic uncertainty, which is taken to be the quadratic difference between the JER values for data and simulation [127]. The energies of the jets are then smeared by this value, allowing the effect in normalisation and shape to be quantified for the discriminating variable in each analysis region. The smearing procedure is one-sided by nature (since smearing cannot increase resolution) meaning that the uncertainty needs to be symmetrised separately.

7.7.5 Jet Reconstruction

Owing to a small discrepancy of 0.2 % between data and simulation in jet reconstruction efficiency for jets with $25\,\text{GeV} < p_\text{T} < 30\,\text{GeV}$, 0.2 % of data jets falling into this category are removed at random, following which all variables which are dependent on jet kinematics are recalculated, and the event selection is repeated.

7.7.6 Heavy- and Light-Flavour Tagging

The efficiency of the b-tagging algorithm is measured separately for the various jet flavours using data-derived control samples which are then compared to simulation. In this way, scale factors may be computed to correct for any discrepancies between the data and simulation.

For b-jets, these scale factors and their corresponding uncertainties are calculated by comparing observed and simulated b-tagging rates in $t\bar{t}$ events [168]. For c-jets, samples containing jets with identified D^* mesons are used [169]. For these heavy flavours, the scale factors are considered to be functions of jet p_T. For light-flavour jets, the scale factors are computed using samples of dijet events, and they are considered as functions of jet p_T and η [170].

There are six independent sources of uncertainty for b-tagging and c-tagging, each of which corresponds to an eigenvector of a matrix containing the total uncertainty per p_T bin and the bin-to-bin correlations. For light jets, there are 12 sources of uncertainty. The various sources of systematic uncertainty are taken to be uncorrelated between b-jets, c-jets, and light jets.

7.7.7 $t\bar{t}$ + jets Modelling

The expected inclusive $t\bar{t}$ cross-section is computed to be 253^{+15}_{-16} pb at $\sqrt{s} = 8\,\text{TeV}$, therefore this cross section is associated with an uncertainty of +5 %/-6 % [143–145]. The shape uncertainties in $t\bar{t}$ modelling arise from the implementation of parton showering and hadronisation in the simulation, in addition to uncertainties related to the reweighting scheme which improves agreement between simulation and data. Additional uncertainties are associated to $t\bar{t}$ production in association with heavy flavours.

The reweighting procedure is applied to $t\bar{t}$ events in the simulation by considering the data-MC difference in p_{T} between the top quark and the $t\bar{t}$ system at $\sqrt{s} = 7\,\text{TeV}$ [171]. The nine most important sources of systematic uncertainty in these p_{T} measurements, together accounting for 95 % of the total, are each considered as separate contributions to the uncertainty in the reweighting factors.

The main contributions to the uncertainty on the measurement of the differential p_{T} distributions are radiation modelling, generator choice, and the uncertainties associated with jet energy resolution, jet energy calibration, and jet flavour tagging. While it could in principle be necessary to consider the correlations between the uncertainties in jet energy calibration at $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 8\,\text{TeV}$, the corresponding difference in expected sensitivity has been found to be extremely low, meaning that it is viable to treat these uncertainties as uncorrelated.

Since this reweighting procedure is performed on an inclusive sample of $t\bar{t}$ events, and thus the uncertainty specifically on the $t\bar{t}$+HF component is unknown, two additional sources of uncertainty are introduced to apply to these events. These uncertainties quantify the difference between applying and not applying the reweighting to the $t\bar{t}$ system p_{T} and the top quark p_{T} respectively.

The uncertainty arising from the choice of parton shower and hadronisation model is calculated by comparing events generated by PowHeg, interfaced with Pythia and Herwig respectively. There are two key sources of uncertainty here: the effect of parton shower choice on the number of jets, and the effect on the heavy flavour content of the jets. Therefore, the parton shower model uncertainty is represented by two uncorrelated parameters – one for $t\bar{t}$+light and one for $t\bar{t}$+HF.

The $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ components are each assigned a conservative normalisation uncertainty of 50 %, corresponding to the upper limit on the theoretical uncertainties in their cross-sections as derived from phenomenology studies [172, 173], and they are

treated as uncorrelated. This approach allows high-statistics LHC data to act as a further constraint on these uncertainties within this analysis.

There are further systematic uncertainties associated to variations in the factorisation and renormalisation scales, the matching threshold and the variations in the masses of the b- and c-quarks in the simulation. The difference between $t\bar{t}$+HF production in MadGraph+Pythia and PowHeg+Pythia is considered to be a separate source of uncertainty.

7.7.8 W/Z+jets Modelling

The W/Z+jets backgrounds are derived from simulation and normalised to their theoretical cross-sections. A reweighting is applied in order to improve the modelling of the W and Z p_{T} spectra, and the absolute difference between applying and not applying this reweighting is taken to be the source of systematic uncertainty. This is then symmetrised.

7.7.9 Multijet Modelling

Multijet events can act as a background due to the potential misidentification of jets as leptons (*i.e.* incorrect event reconstruction). For both e+jets and μ+jets, the multijet background contains non-prompt leptons (such as muons originating from hadron decays), and in the e+jets channel there is also a significant background contribution due to "fake" electrons – electrons from photon conversions within the electromagnetic calorimeter, and jets that deposit large amounts of energy into the electromagnetic calorimeter and are incorrectly reconstructed as electrons.

Limited data statistics mean that the systematic uncertainties on the multijet backgrounds are not well constrained; this is especially true in the high-jet-multiplicity and high-b-tag-multiplicity categories. There are further contributions from the lepton fake rates. These uncertainties are taken to be uncorrelated between the different jet and b-tag multiplicity categories. An uncertainty of 50 % is assigned to the lepton fake rate, which is taken as correlated across the various jet and b-tag multiplicity categories, but uncorrelated between the electron and muon channels.

7.7.10 Electroweak Background Modelling

The theoretical uncertainty on the single-top production cross-section is taken to be +5 %/-4 %. This is computed by taking the weighted average of the theoretical uncertainties on single-top production in the s, t, and Wt channels [174, 175].

The uncertainty on diboson rates is calculated as the sum in quadrature of the theoretical cross-section uncertainty of ±5 % with an uncertainty of ±24 %, relating to the extrapolation into high-jet-multiplicity bins [139].

The theoretical cross-section uncertainty on $t\bar{t}V$ production is taken to be ±30 % – this is the only source of uncertainty in the $t\bar{t}V$ background which was found to have a non-negligible effect [176, 177].

7.7.11 Signal Modelling

The effect of factorisation and renormalisation scale choice on the signal kinematics is quantified by using NLO POWHEL $t\bar{t}H$ samples, within which the default scale value is varied up and down by a factor of two. A reweighting scheme was used to introduce the observed effect of these variations into the nominal $t\bar{t}H$ sample. Similarly, a reweighting was used to reproduce the effect of changing between static and dynamic scales. Systematic uncertainties arising from parton shower and fragmentation choice, or choice of PDF set, were expected to have a small effect on the final result and they were therefore not considered. Information about the signal cross-section uncertainty is taken from Ref. [37].

7.8 Standard Model Analysis

Within each of the analysis regions, the objective was to understand how to maximise the separation between the semileptonic $t\bar{t}H, H \to b\bar{b}$ signal (for the mass-point $m_H = 125\,\text{GeV}$) and the given background processes. The main contributions to the standard model background are $t\bar{t}+\text{HF}$ (heavy flavours, *i.e.* b- and c-quarks) and $t\bar{t}+nj$ (light jets, which can be misidentified as b-jets). The main function of the TMVA training procedures in the signal regions is to find a variable which has optimal discrimination power with respect to these backgrounds.

7.8.1 Analysis Method

The multilayer perceptron neural network, implemented in the TMVA within the GlaNtp framework, is trained in those regions where a significant contribution from $t\bar{t}H, H \to b\bar{b}$ signal events is anticipated. Following the labelling scheme desribed above, these regions are (5,4) (6,3) and (6,4). In signal-depleted regions (where the neural network has less classification power) the single variable $H_{\mathrm{T}}^{\mathrm{had}}$ (*i.e.* the scalar sum of all jet p_{T}) is used as a discriminating variable. These TMVA score distributions are then used as inputs to code which performs a combined fit to data, including the six control regions.

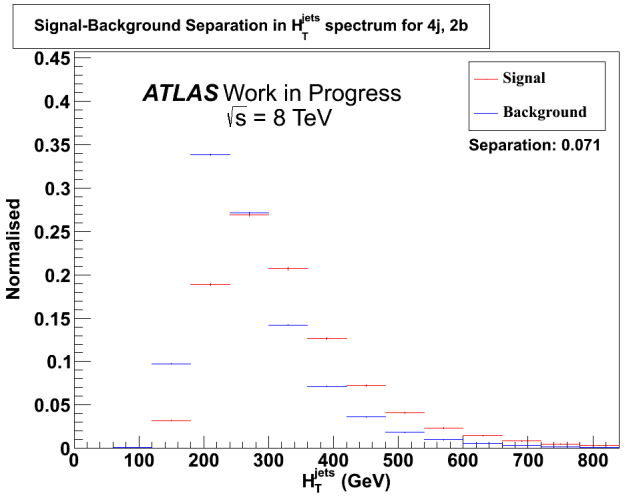
7.8.2 Neural Network Input Variables

Table 7.3 lists the input variables on which the TMVA was trained, together with their definitions. From this list an appropriate set of variables is chosen for each of the four analysis regions in which neural network training is performed. This set represents a wide selection of variables – including kinematic, event-shape and object-pair variables – which have been chosen for their potential (dependent on analysis region) to increase sensitivity. The relative importance (*i.e.* relative statistical separation power) of each variable in a given instance of neural network training will be computed by the TMVA and displayed in its text output to enable comparisons. In each analysis region, a set comprising seven to ten of the most highly ranked input variables is chosen so as to maximise the TMVA's sensitivity; including more variables will typically yield decreasing returns.

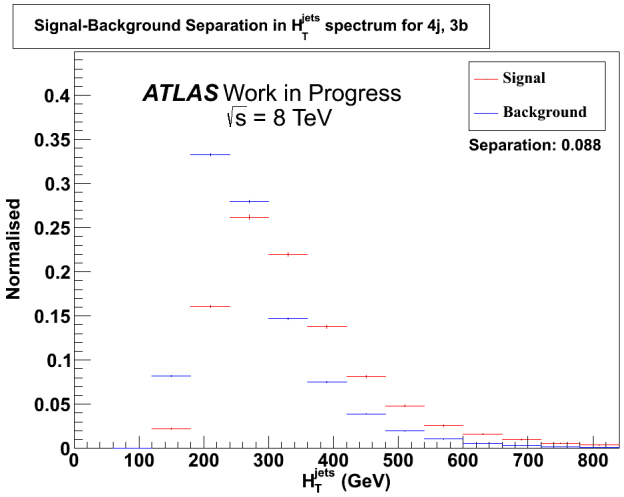
Note that the kinematic variables in those regions with ≥ 6 jets are constructed by considering the highest-ranked seven jets – ranking all b-tagged jets by p_{T} and then all other jets, again ranked by p_{T}. It is furthermore required that all training variables (together with the correlations between them) be well-described by data in several control regions.

Variable	Definition
H_0	Fox-Wolfram moment of order zero, computed using all jets and the lepton
H_1	Fox-Wolfram moment of order one, computed using all jets and the lepton
$H_{\mathrm{T}}^{\mathrm{had}}$	Scalar sum of all jet p_{T}
aplanarity$_{b-\mathrm{jet}}$	Equal to $1.5\lambda_2$ (λ_2 is the second eigenvalue of the momentum tensor constructed using only b-jets)
centrality	Centrality = $\Sigma p_{\mathrm{T}}/\Sigma E$ (summing over all jets and the lepton)
$m_{\mathrm{bj}}^{\mathrm{max}p_{\mathrm{T}}}$	Invariant mass of the pair comprising a b-jet and any jet, chosen so as to maximise pair p_{T}
$m_{\mathrm{jj}}^{\mathrm{max}p_{\mathrm{T}}}$	Invariant mass of the pair comprising any two jets, chosen so as to maximise pair p_{T}
$m_{\mathrm{bj}}^{\mathrm{min}\Delta R}$	Invariant mass of the pair comprising a b-jet and any jet, chosen so as to minimise pair ΔR
$m_{bb}^{\mathrm{min}\Delta R}$	Invariant mass of the pair comprising two b-jets, chosen so as to minimise pair ΔR
$m_{jj}^{\mathrm{min}\Delta R}$	Mass of the jet-jet pair with minimum ΔR
$m_{jj}^{\mathrm{max}p_{\mathrm{T}}}$	Mass of the jet-jet pair with maximum p_{T}
$m_{uu}^{\mathrm{min}\Delta R}$	Mass of the pair of untagged jets, chosen so as to minimise ΔR
m_{jjj}	Mass of the jet triplet with the largest vector-sum p_{T}
$m_{W_{\mathrm{had}}}$	Reconstructed mass of the hadronic W boson
$\Delta R_{\mathrm{lep}-bb}^{\mathrm{min}\Delta R}$	ΔR of the lepton and the pair of b-jets with minimum ΔR
$\Delta R_{bb}^{\mathrm{max}p_{\mathrm{T}}}$	ΔR of the pair of b-jets with maximum p_{T}
$\Delta R_{bb}^{\mathrm{avg}}$	Averaged ΔR of all bb pairs
N_{40}^{jet}	Number of jets in event satisfying $p_{\mathrm{T}} > 40\,\mathrm{GeV}$
$p_{\mathrm{T}}^{\mathrm{jet5}}$	p_{T} of the fifth-leading jet
$p_{\mathrm{T}}^{jj,\mathrm{av}}$	Average p_{T} of a jet-jet pair

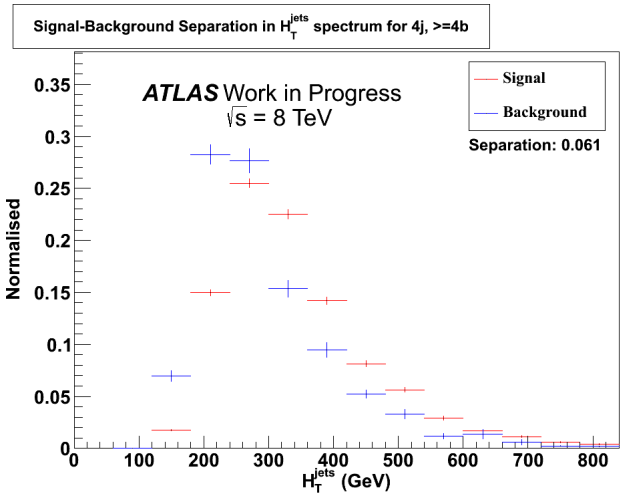
Table 7.3: Table showing all variables which are used as inputs to TMVA training in any region. The Fox-Wolfram moments H_l are defined in Ref. [178].



(a) (4,2)

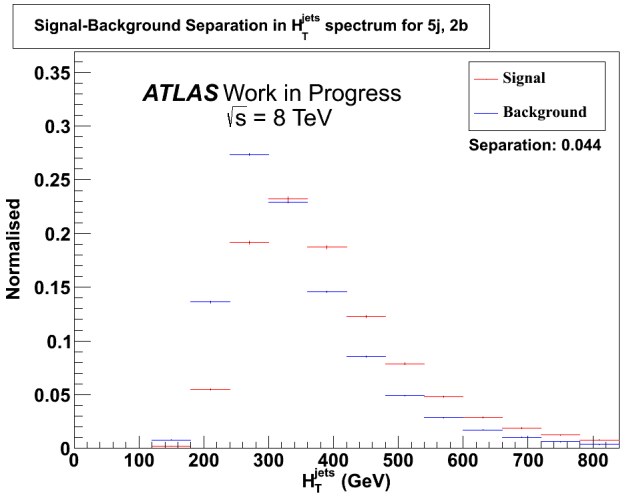


(b) (4,3)

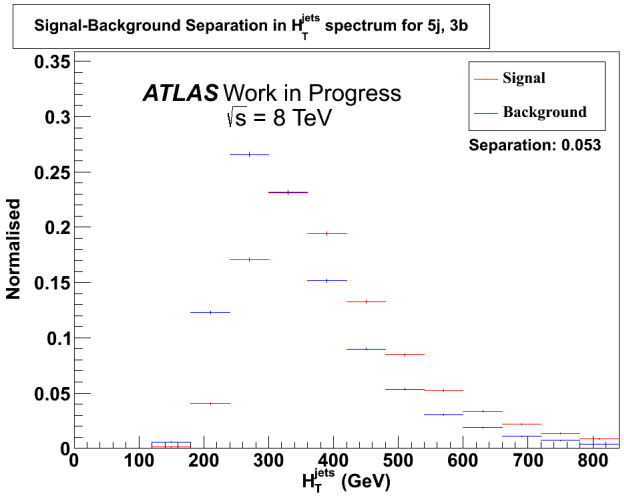


(c) (4,4)

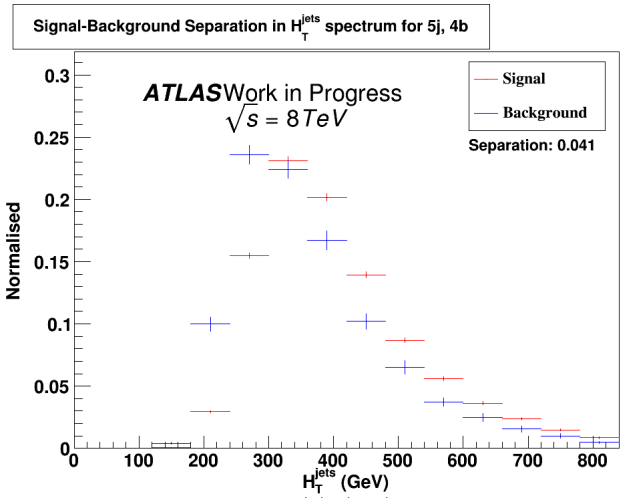
Figure 7.2: Plots showing the shape difference and separation in the $H_{\mathrm{T}}^{\mathrm{had}}$ spectra for the analysis regions (4,2), (4,3), and (4,4).



(a) (5,2)

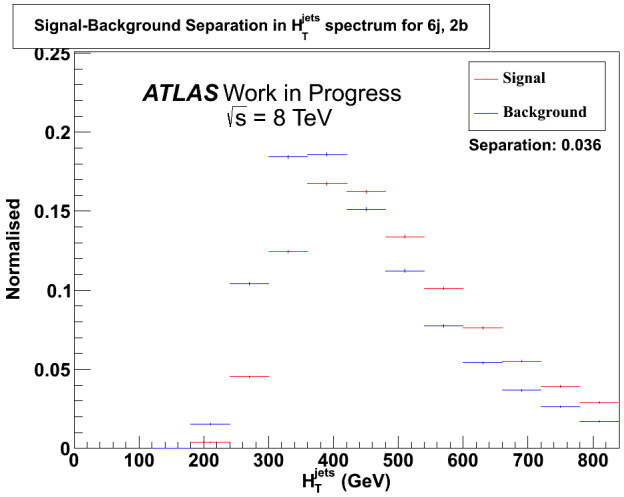


(b) (5,3)

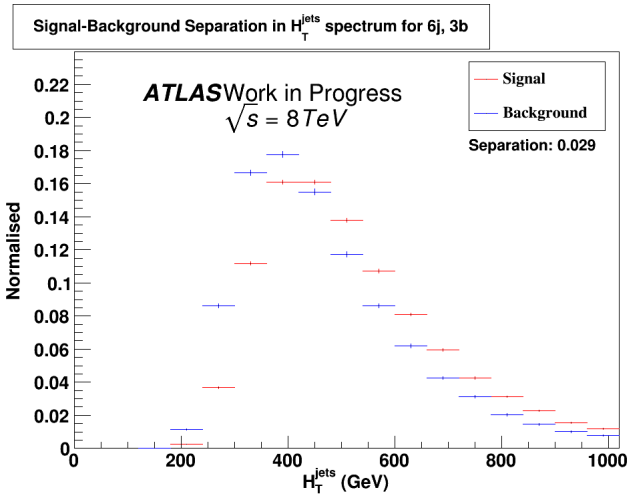


(c) (5,4)

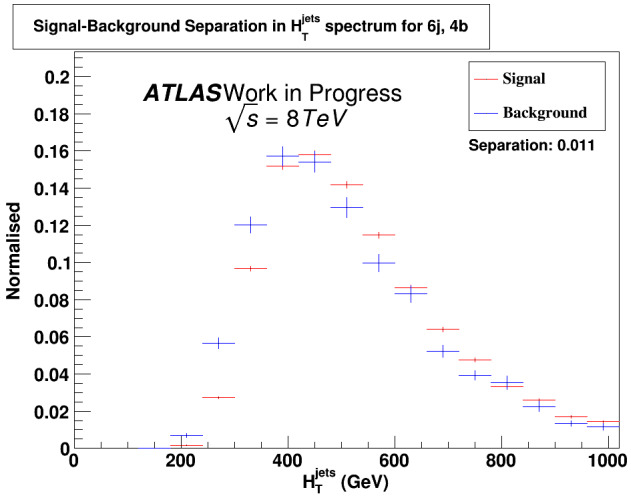
Figure 7.3: Plots showing the shape difference and separation in the $H_{\mathrm{T}}^{\mathrm{had}}$ spectra for the analysis regions (5,2), (5,3), and (5,4).



(a) (6,2)



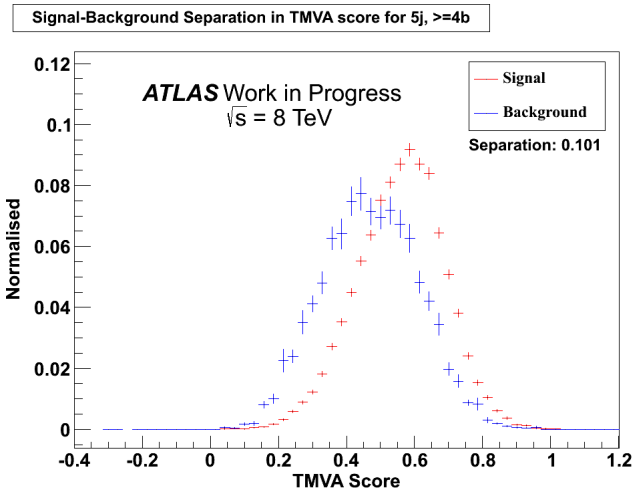
(b) (6,3)



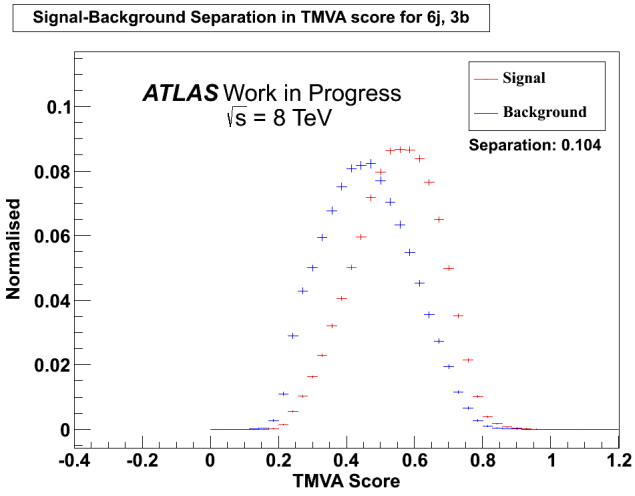
(c) (6,4)

Figure 7.4: Plots showing the shape difference and separation in the $H_{\mathrm{T}}^{\mathrm{had}}$ spectra for the analysis regions (6,2), (6,3), and (6,4)

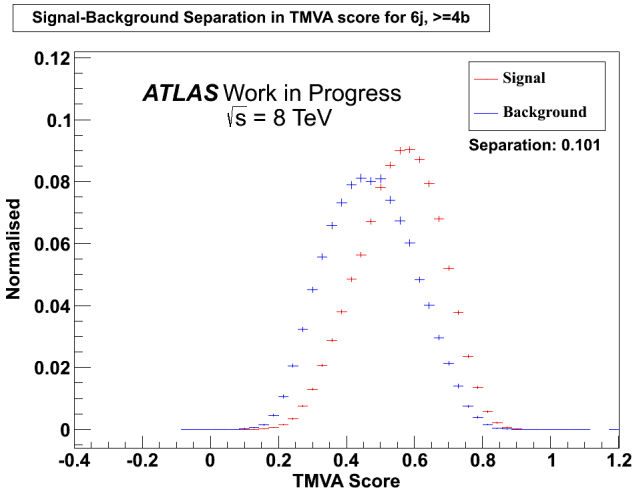
.



(a) (5,4)



(b) (6,3)



(c) (6,4)

Figure 7.5: Plots showing the shape difference and separation in the TMVA score distributions for the analysis regions (5,4), (6,3), and (6,4).

	Region (5,4)			Region (6,3)			Region (6,4)		
Rank	Variable	Importance	Separation	Variable	Importance	Separation	Variable	Importance	Separation
1	aplanarity$_{b-\text{jet}}$	45	1.75	H_1	12	0.76	H_0	74	1.41
2	H_1	3.7	3.37	centrality	3.4	1.96	aplanarity$_{b-\text{jet}}$	54	1.31
3	$\Delta R_{bb}^{\text{avg}}$	2.0	3.72	$\Delta R_{bb}^{\text{avg}}$	0.66	1.41	H_1	45	2.12
4	centrality	1.8	2.22	$p_{\text{T}}^{\text{jet5}}$	0.38	1.44	centrality	3.9	1.58
5	$H_{\text{T}}^{\text{had}}$	0.19	1.73	$\Delta R_{\text{lep}-bb}^{\text{min}\Delta R}$	0.098	0.74	$m_{bj}^{\text{max}p_{\text{T}}}$	0.33	1.81
6	N_{40}^{jet}	0.17	0.18	N_{40}^{jet}	0.064	0.95	$p_{\text{T}}^{\text{jet5}}$	0.27	1.02
7	$\Delta R_{\text{lep}-bb}^{\text{min}\Delta R}$	0.023	1.61	$m_{bj}^{\text{min}\Delta R}$	0.011	0.46	$H_{\text{T}}^{\text{had}}$	0.15	1.01
8	$m_{bb}^{\text{min}\Delta R}$	0.014	2.39	$m_{uu}^{\text{min}\Delta R}$	0.030	0.26	N_{40}^{jet}	0.15	0.04
9	$m_{jj}^{\text{min}\Delta R}$	0.012	1.04	$m_{bb}^{\text{min}\Delta R}$	0.0023	0.20	$p_{\text{T}}^{jj,\text{av}}$	0.070	1.23
10	$m_{jj}^{\text{max}p_{\text{T}}}$	0.0015	2.37	$m_{bj}^{\text{max}p_{\text{T}}}$	0.00080	0.52	m_{jjj}	0.0028	1.29

Table 7.4: The ten highest-ranked TMVA input variables in each of the three signal-enriched regions. Separation is as defined in Equation (7.1) and importance is as defined in Equation (7.2). The table is ordered by importance, which is rounded to two significant figures.

7.8.3 TMVA Training by Region

Table 7.3 details the set of variables over which the neural network was trained. A subset of these was chosen, according to their respective rankings and their correlations with one another, to be used in each training region. It is desirable to train the neural network on a set of input variables which is diverse (minimal linear dependency, *i.e.* minimal redundancy) and which individually give good signal-background separation. This information is presented by TMVA after training is complete, allowing the training to be optimised iteratively.

Figures 7.2, 7.3, and 7.4 show the signal-background separation in the H_{T} spectra for all analysis regions. The signal-background separations achieved by TMVA training in the analyis regions are shown in Figure 7.5. In each of these plots, for TMVA score and H_{T}, the separation $\langle S^2 \rangle$ of a classifying variable y is computed according to [114]:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_{\mathrm{S}}(y) - \hat{y}_{\mathrm{B}}(y))^2}{\hat{y}_{\mathrm{S}}(y) + \hat{y}_{\mathrm{B}}(y)} \, \mathrm{d}y, \tag{7.1}$$

where $\hat{y}_{\mathrm{S}}(y)$ represents the signal PDF and $\hat{y}_{\mathrm{B}}(y)$ the background PDF. This integral gives a scalar value which quantifies how distinct the signal and background distributions are – intuitively, the sensitivity of the analysis depends strongly on this separation value, especially in the signal-enriched regions (within which the separation measures the discriminating power of the neural network). $\langle S^2 \rangle = 1$ corresponds to perfect separation, *i.e.* non-overlapping signal and background distributions, while $\langle S^2 \rangle = 0$ corresponds to full overlap, *i.e.* indistinct signal and background distributions. $\langle S^2 \rangle$ is used to examine and compare the separation power of different variables used in this analysis.

For a multilayer perceptron training, the TMVA internally ranks its inputs by a metric related to the sum of the squared weights of the connections between the nodes in the input layer (corresponding to each respective input) and the nodes in the first hidden layer. The importance I_i of input variable x_i is defined as

$$I_i = \bar{x_i}^2 \sum_{j=1}^{n_{\mathrm{h}}} (w_{ij})^2 \, , \tag{7.2}$$

where w_{ij} denotes the weight associated to the connection between input node i and hidden node j, and the sum runs over the n_{h} nodes in the first hidden layer [114]. Table 7.4 shows the importance and separation for each of the selected input

variables in each analysis region for which neural network training was performed. GlaNtp also includes functionality to allow it to plot the input variables to neural network training. Stacked plots of the three highest-ranked variables (by TMVA importance) are shown in Appendix C.

7.8.4 Results

A combined fit was performed over the spectra of the discriminating variables for each of the nine analysis regions – TMVA score in those regions where neural network training was performed, and $H_{\mathrm{T}}^{\mathrm{had}}$ in the control regions. No further cut was imposed on these variables prior to performing the fit to data. A likelihood function $\mathcal{L}(\mu, \theta)$ (where μ is the fitted Higgs signal-strength parameter and θ denotes the set of nuisance parameters) determined the expected number of events in each bin; μ was required to be the same across all analysis regions. The nuisance parameters are a statistical representation of the effects of the systematic uncertainties on the content of each bin; the total likelihood may be considered as a product of these Poisson bin-by-bin fluctuations. The effect of the systematics was constrained by including the control regions in the fit; these may be considered as high-statistics, overwhelmingly background-dominated regions, with a small contamination of signal events.

Figures 7.6–7.9 display the respective pulls of each source of systematic uncertainty (as defined in Equation (6.14)) following the fits to Asimov data and to real data under the signal-plus-background hypothesis. Figures 7.10, 7.11, and 7.12 display the contributions of each physical process to each bin in the analysis, preceding and following the fit to data. In each of the signal-depleted regions, eleven bins are used for the $H_{\mathrm{T}}^{\mathrm{had}}$ spectra, while thirteen bins are used for the neural network output in the signal-enriched regions.

Table 7.6 shows the pre-fit yields of events in each region, following MVA event selection. For each process, the uncertainty is computed as the sum in quadrature of each individual uncertainty contribution. These yields differ somewhat from the median prefit yields observed by the ATLAS $t\bar{t}H$ working group, by a factor of typically around 15–20 % [93]. Specifically, the yields shown here are consistently lower, for both data and Monte Carlo, consistent with GlaNtp applying a more stringent event selection, as documented in Appendix B. The corresponding post-fit yields are shown in Table 7.7. The effective scaling applied between the prefit and postfit yields for each process is shown in Table 7.8. It is notable that in each region, the multijet background is significantly scaled down by the fit to data, and

Region	Expected S/B (%)	Expected $S/\sqrt{B}$
(4,2)	0.03	0.10
(4,3)	0.16	0.13
(4,4)	1.71	0.19
(5,2)	0.09	0.18
(5,3)	0.40	0.28
(5,4)	2.37	0.34
(6,2)	0.25	0.37
(6,3)	0.97	0.57
(6,4)	3.78	0.73

Table 7.5: Table showing the expected signal-to-background and signal-to-root-background ratios for each analysis region, computed using the prefit yields as shown in Table 7.6.

the $t\bar{t}H$ signal is scaled up, while remaining a minor contribution to the overall postfit yields observed in each analysis region. Moreover, the irreducible heavy-flavour backgrounds (especially $t\bar{t}c\bar{c}$), as well as $t\bar{t}$+light jets, are scaled up, while the event yields of the reducible backgrounds such as $t\bar{t}+V$, single-top production, and V+jets are not greatly changed by the fit in most categories. The expected ratios of signal to background (S/B) and signal to root background ($S/\sqrt{B}$) are shown in Table 7.5.

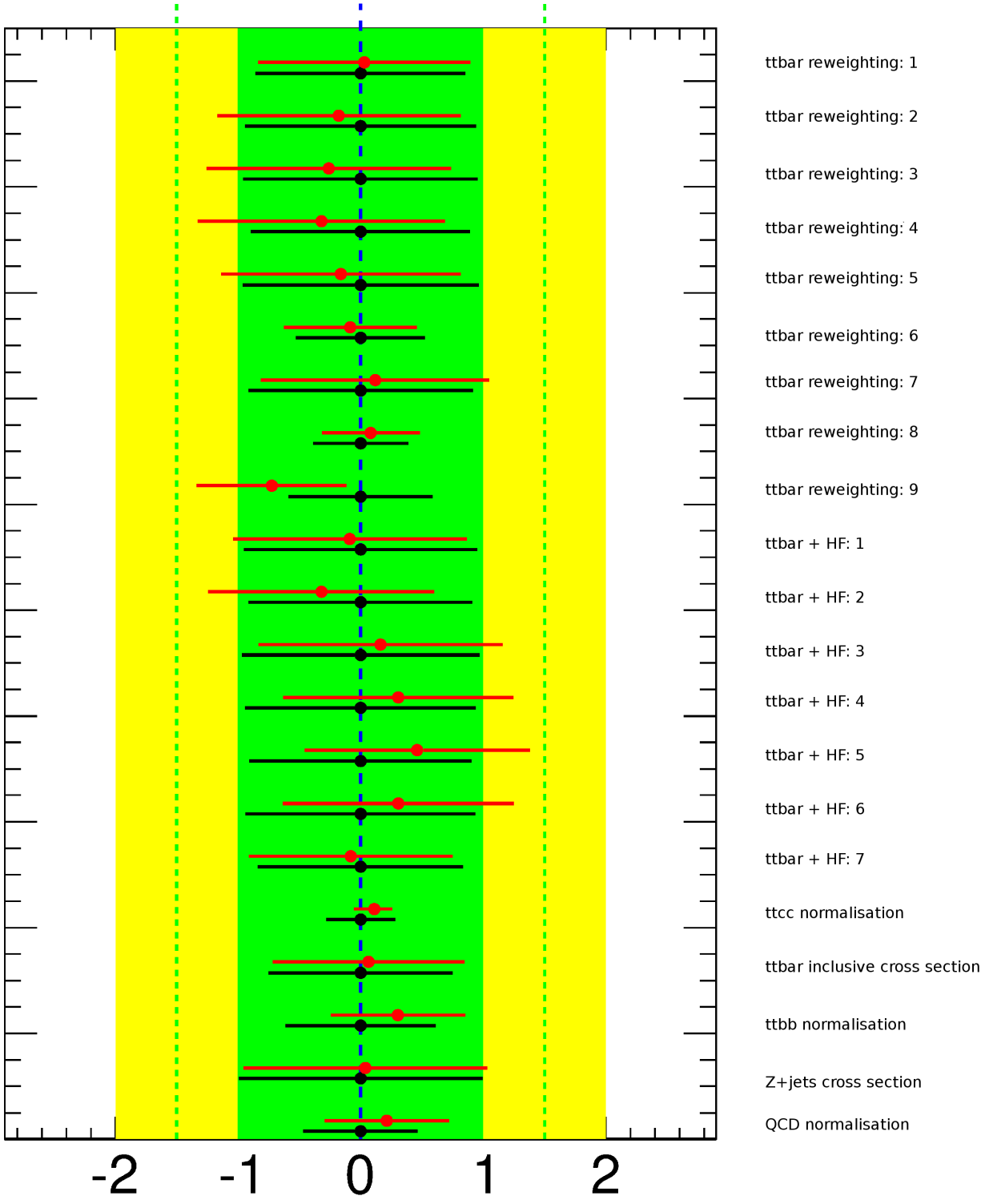


Figure 7.6: Nuisance parameters fitted to Asimov data (black) and data (red) across all analysis regions for the signal-plus-background hypothesis. Numbers identify components of a given systematic uncertainty; for tagging parameters, these numbers denote the rank of that eigenvector component of the mistag matrix; for other parameters they are abbreviations which are detailed in Appendix D. Here, $t\bar{t}$ modelling parameters are shown.

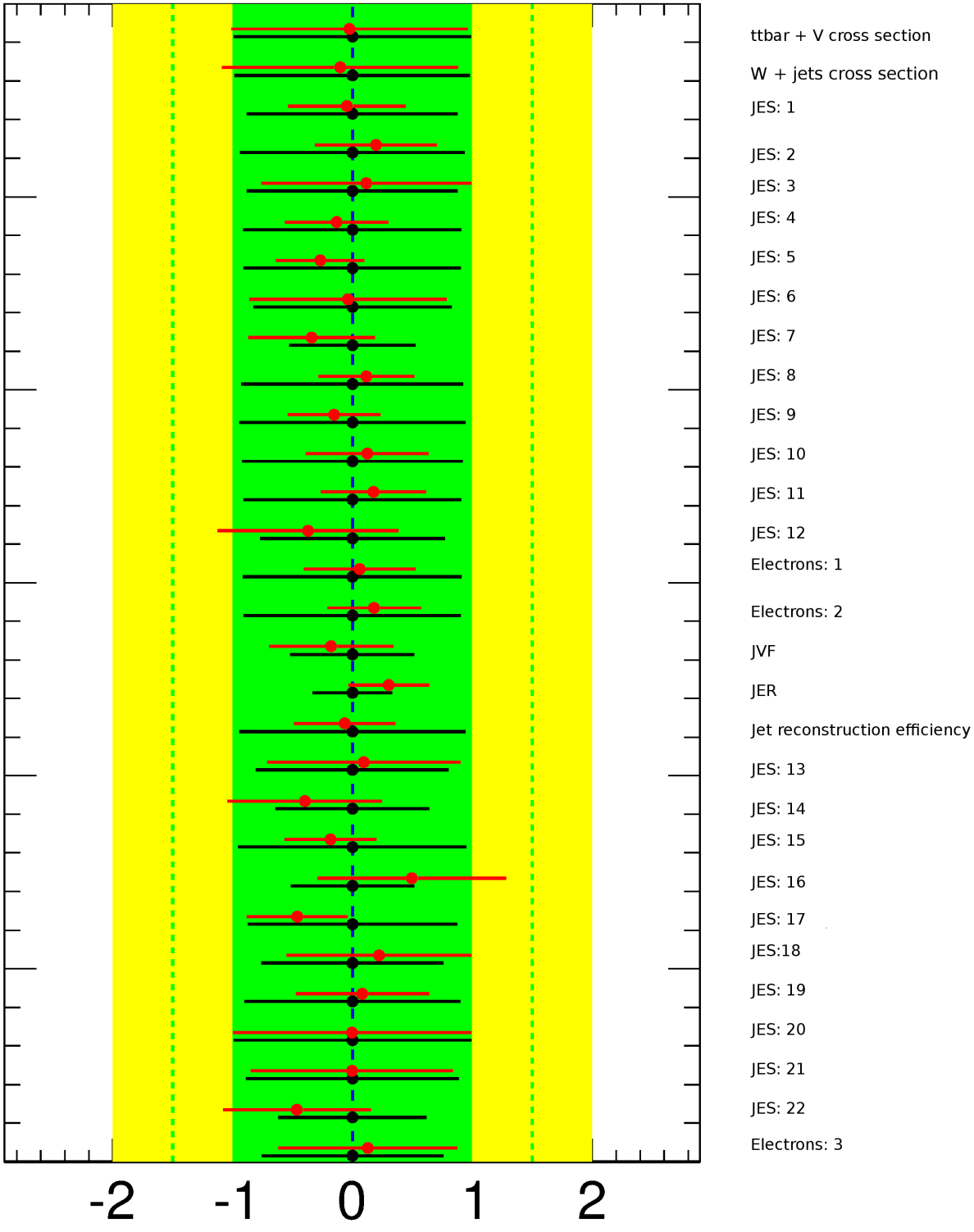


Figure 7.7: Nuisance parameters fitted to Asimov data (black) and data (red) across all analysis regions for the signal-plus-background hypothesis. Numbers identify components of a given systematic uncertainty; for tagging parameters, these numbers denote the rank of that eigenvector component of the mistag matrix; for other parameters they are abbreviations which are detailed in Appendix D. Here, jet parameters are shown.

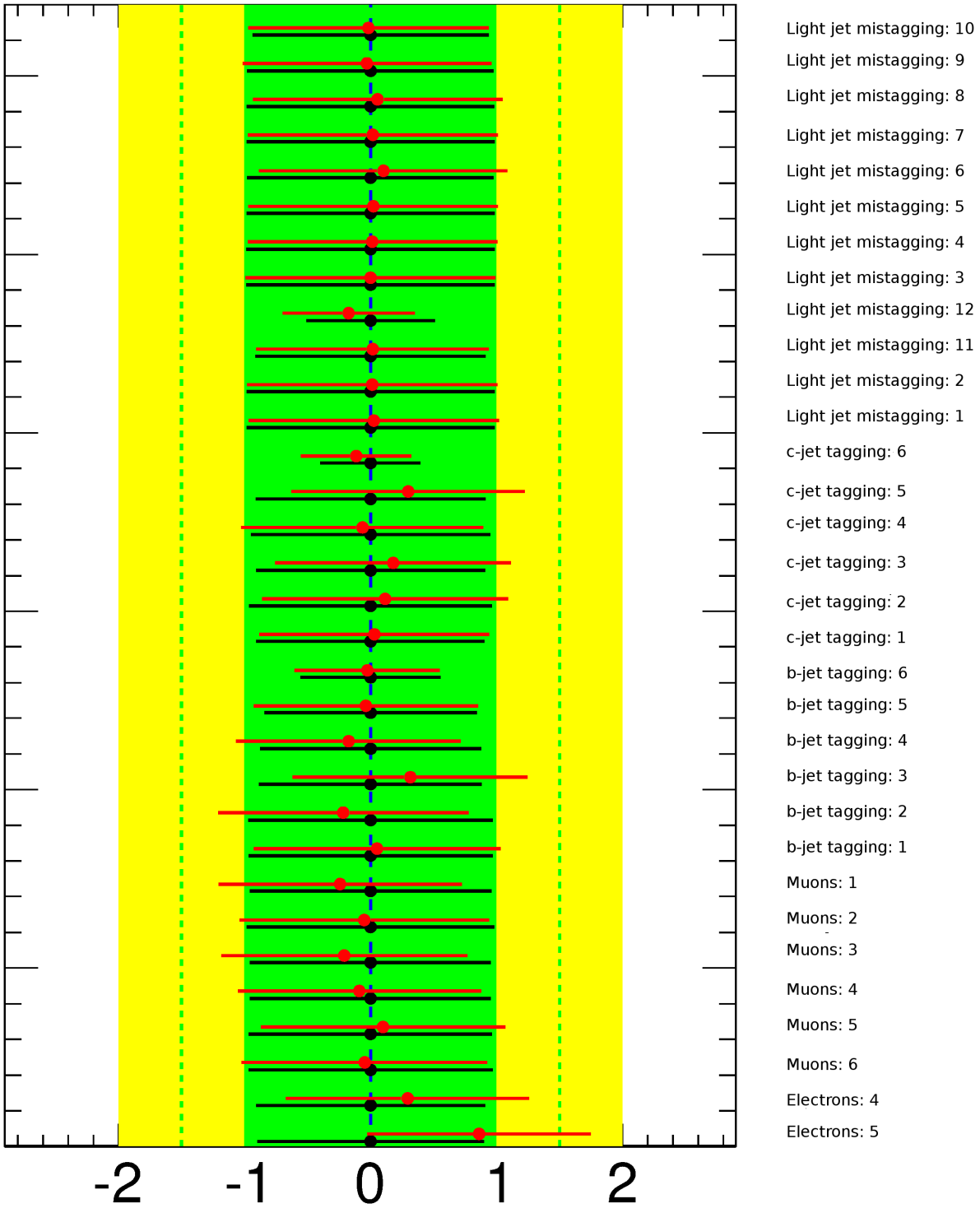


Figure 7.8: Nuisance parameters fitted to Asimov data (black) and data (red) across all analysis regions for the signal-plus-background hypothesis. Numbers identify components of a given systematic uncertainty; for tagging parameters, these numbers denote the rank of that eigenvector component of the mistag matrix; for other parameters they are abbreviations which are detailed in Appendix D. Here, lepton and tagging parameters are shown.

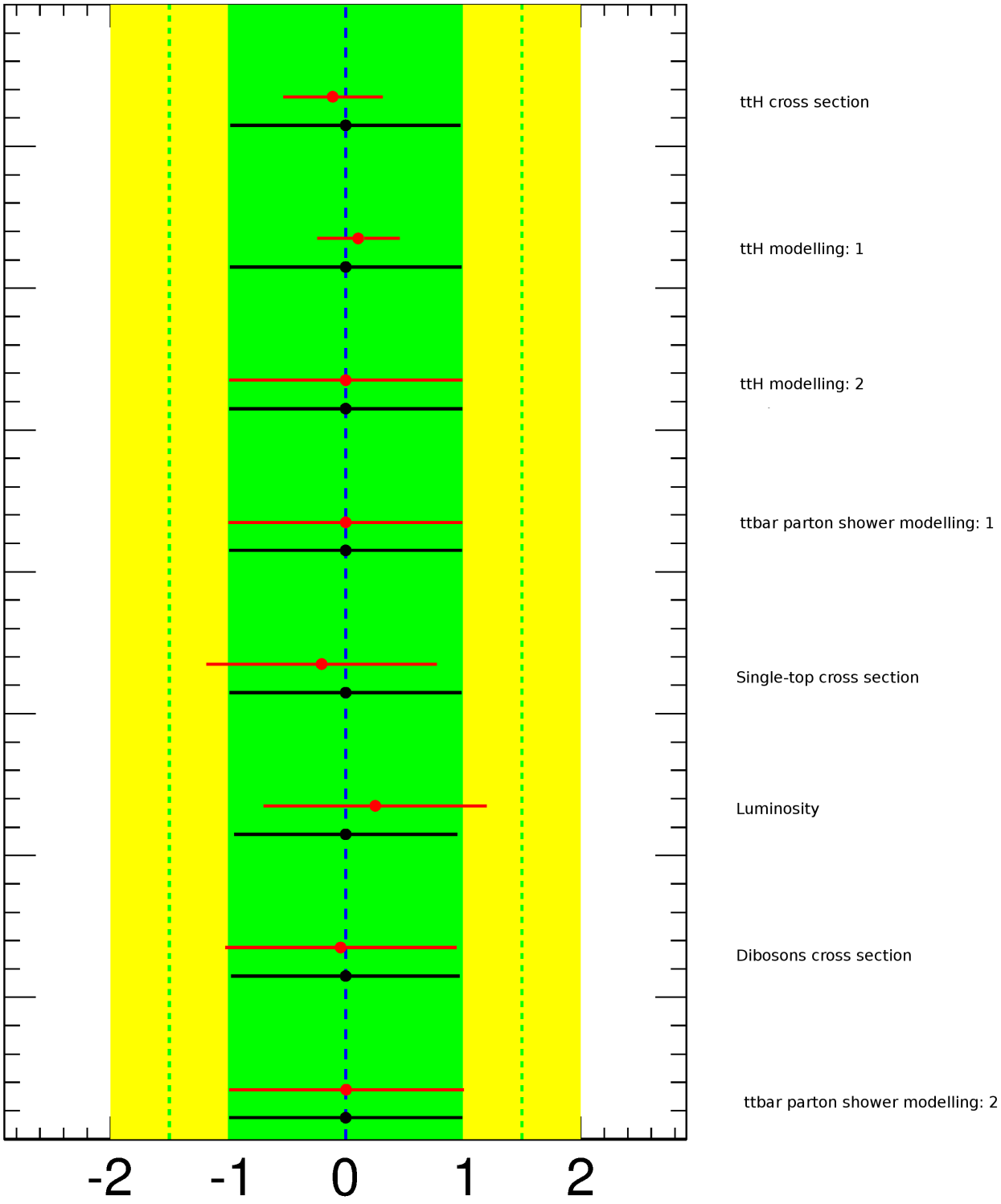
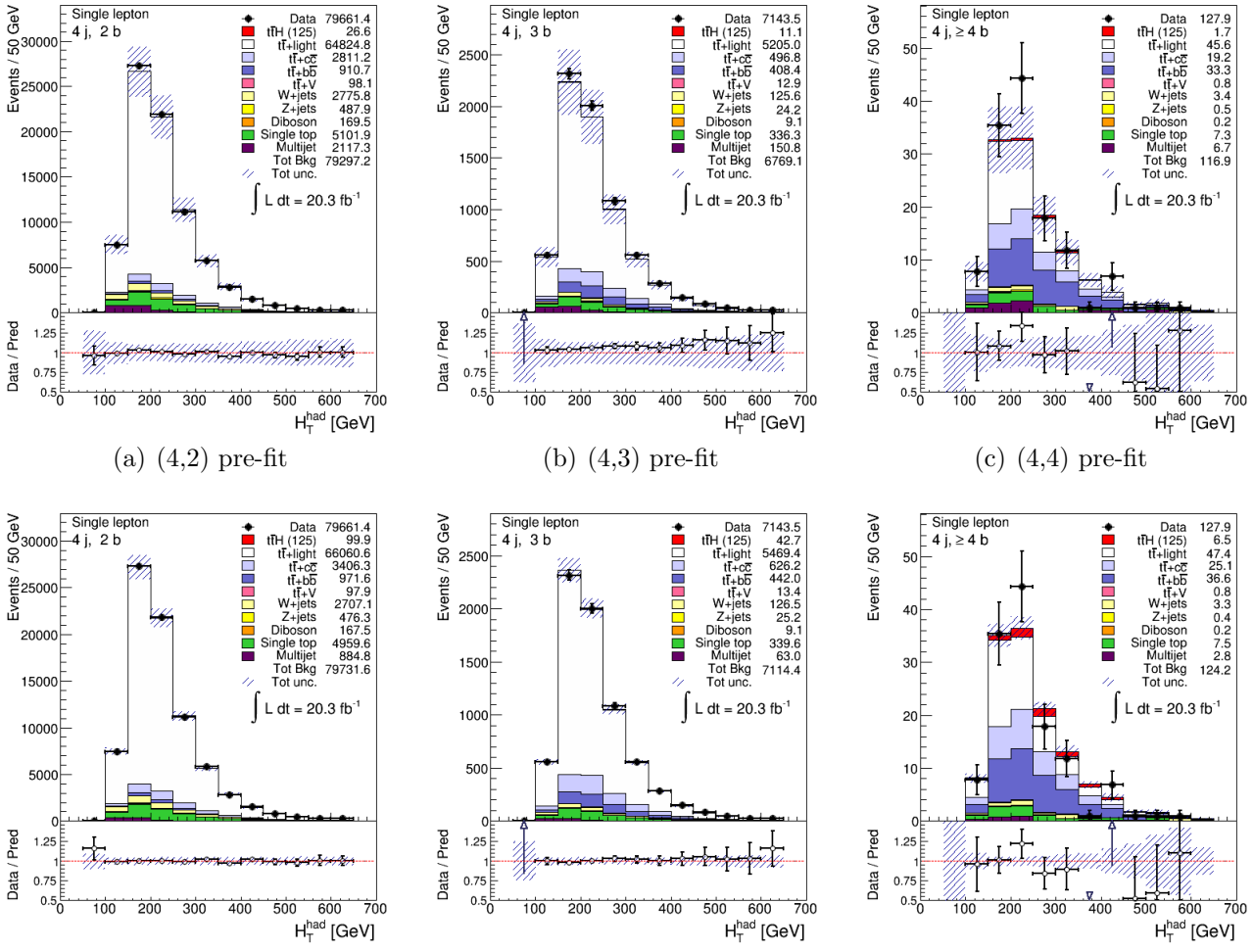


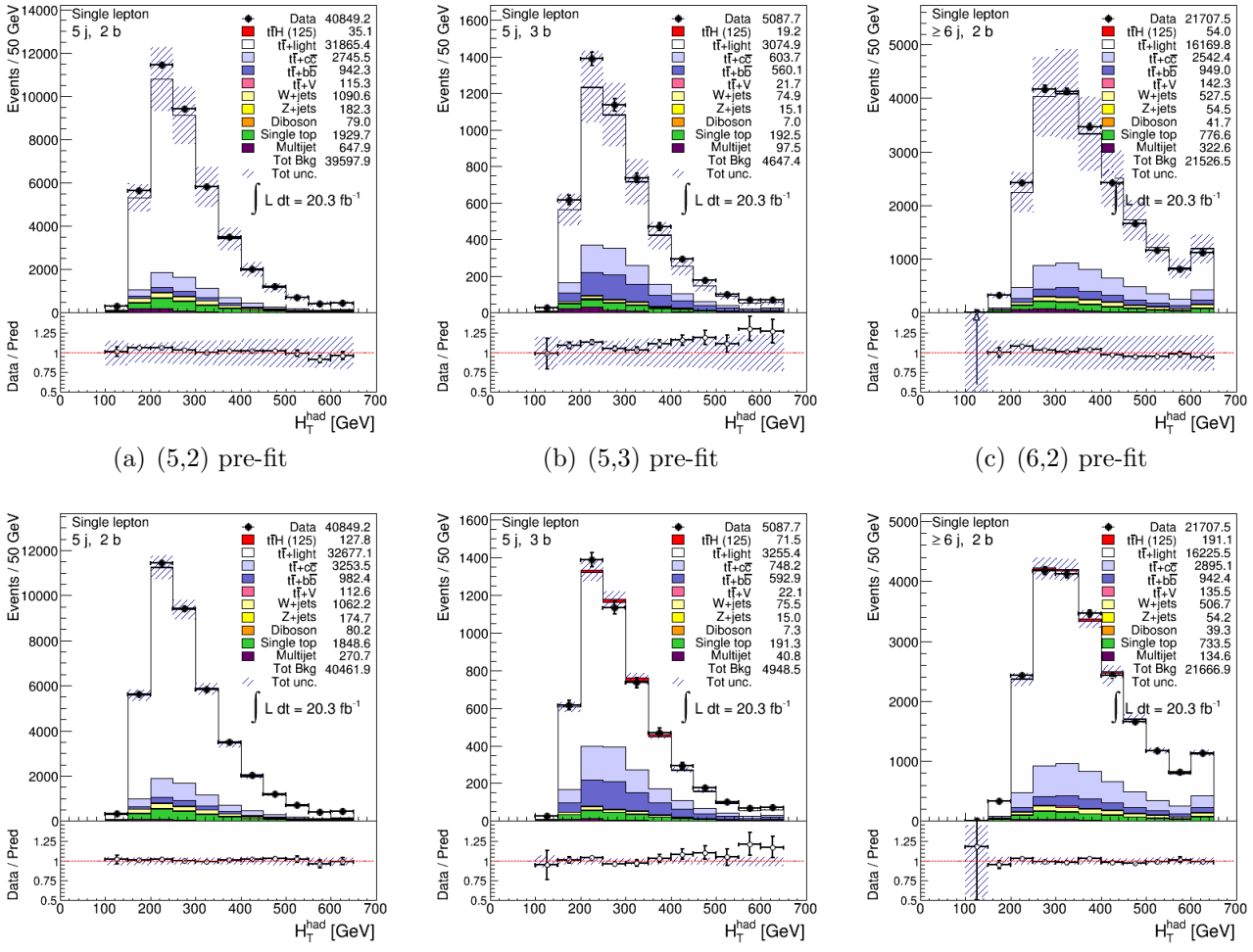
Figure 7.9: Nuisance parameters fitted to Asimov data (black) and data (red) across all analysis regions for the signal-plus-background hypothesis. Numbers identify components of a given systematic uncertainty; for tagging parameters, these numbers denote the rank of that eigenvector component of the mistag matrix; for other parameters they are abbreviations which are detailed in Appendix D. Here, miscellaneous parameters are shown.



(a) (4,2) pre-fit (b) (4,3) pre-fit (c) (4,4) pre-fit

(d) (4,2) post-fit (e) (4,3) post-fit (f) (4,4) post-fit

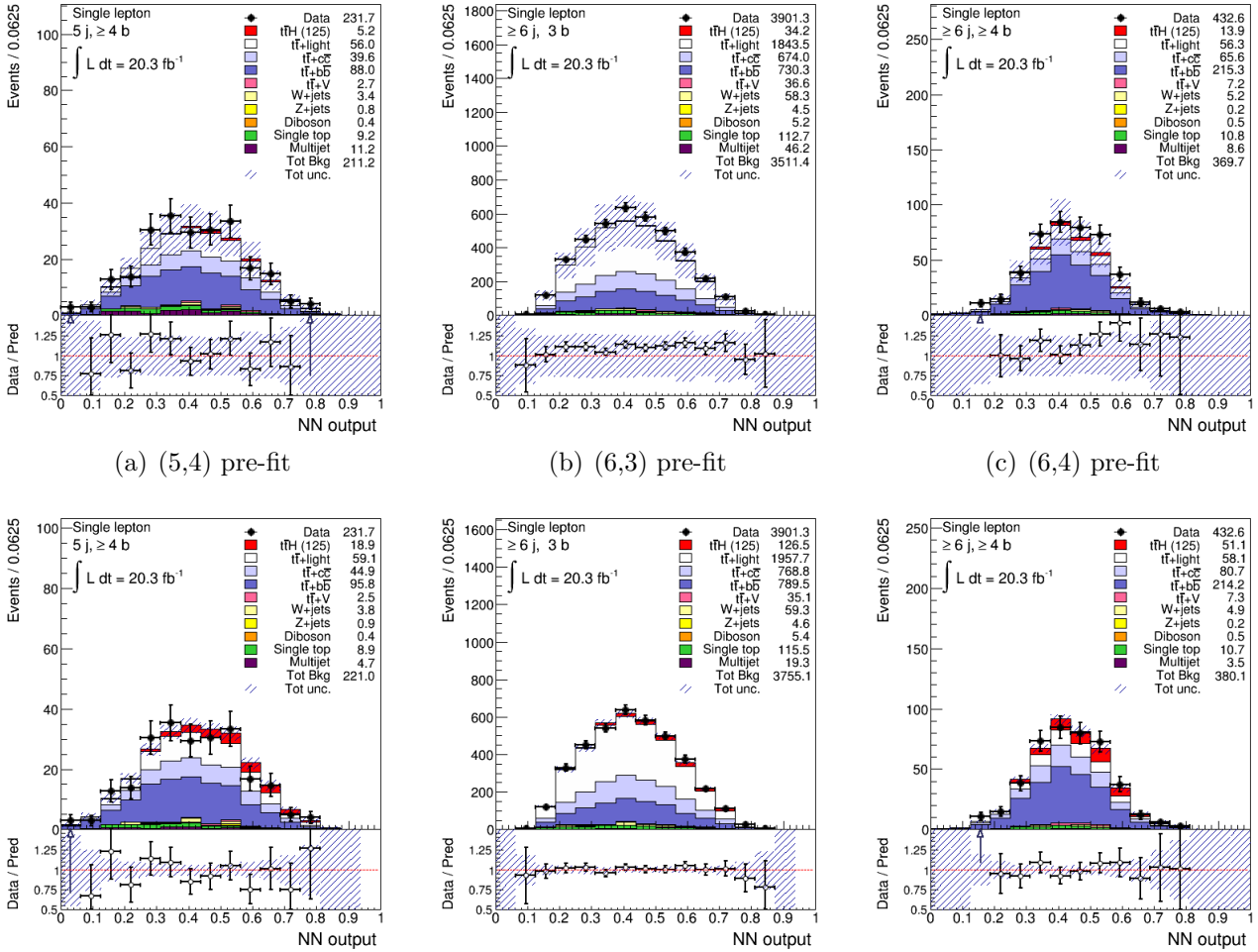
Figure 7.10: Plots showing the stacked contributions to each bin in each of the signal-depleted analysis regions (4,2), (4,3), and (4,4), before and after the fit to data. The numbers in the plot legends indicate the yields of the different processes, in accordance with Tables 7.6 and 7.7.



(a) (5,2) pre-fit (b) (5,3) pre-fit (c) (6,2) pre-fit

(d) (5,2) post-fit (e) (5,3) post-fit (f) (6,2) post-fit

Figure 7.11: Plots showing the stacked contributions to each bin in each of the signal-depleted analysis regions (5,2), (5,3), and (6,2), before and after the fit to data. The numbers in the plot legends indicate the yields of the different processes, in accordance with Tables 7.6 and 7.7.



(a) (5,4) pre-fit (b) (6,3) pre-fit (c) (6,4) pre-fit

(d) (5,4) post-fit (e) (6,3) post-fit (f) (6,4) post-fit

Figure 7.12: Plots showing the stacked contributions to each bin in each of the signal-enriched analysis regions (5,4), (6,3), and (6,4), before and after the fit to data. The numbers in the plot legends indicate the yields of the different processes, in accordance with Tables 7.6 and 7.7.

Region	(4,2)	(4,3)	(4,4)	(5,2)	(5,3)	(5,4)	(6,2)	(6,3)	(6,4)
Multijet	2090 ± 46	150 ± 12	7 ± 3	648 ± 25	97 ± 10	11 ± 3	323 ± 18	46 ± 7	9 ± 3
Single top	5096 ± 71	336 ± 18	7 ± 3	1930 ± 44	193 ± 14	9 ± 3	777 ± 28	113 ± 11	11 ± 3
Diboson	169 ± 13	9 ± 3	0 ± 0	79 ± 9	7 ± 3	0 ± 1	42 ± 6	5 ± 2	1 ± 1
Z + jets	485 ± 22	24 ± 5	0 ± 1	182 ± 13	15 ± 4	1 ± 1	54 ± 7	5 ± 2	0 ± 0
W + jets	2763 ± 53	125 ± 11	3 ± 2	1091 ± 33	75 ± 9	3 ± 2	527 ± 23	58 ± 8	5 ± 2
$t\bar{t} + V$	98 ± 10	13 ± 4	1 ± 1	115 ± 11	22 ± 5	3 ± 2	142 ± 12	37 ± 6	7 ± 3
$t\bar{t} + b\bar{b}$	910 ± 30	408 ± 20	33 ± 6	942 ± 31	560 ± 24	88 ± 9	949 ± 31	730 ± 27	215 ± 15
$t\bar{t} + c\bar{c}$	2811 ± 53	497 ± 22	19 ± 4	2745 ± 52	604 ± 25	40 ± 6	2542 ± 50	674 ± 26	66 ± 8
$t\bar{t}$ + light	64 809 ± 254	5204 ± 72	46 ± 7	31 865 ± 179	3075 ± 55	56 ± 7	16 169 ± 127	1844 ± 43	56 ± 7
Total background	79 231 ± 281	6767 ± 82	117 ± 11	39 598 ± 199	4647 ± 68	211 ± 15	21 525 ± 147	3511 ± 59	370 ± 19
$t\bar{t}H$ signal	27 ± 5	11 ± 3	2 ± 1	35 ± 6	19 ± 4	5 ± 2	54 ± 7	34 ± 6	14 ± 4
Data	79 598	7140	128	40 849	5088	232	21 705	3901	433

Table 7.6: Table showing the yields of each process in each analysis region before the fit to data, rounded to the nearest whole event.

Region	(4,2)	(4,3)	(4,4)	(5,2)	(5,3)	(5,4)	(6,2)	(6,3)	(6,4)
Multijet	873 ± 30	63 ± 8	3 ± 2	271 ± 16	41 ± 6	5 ± 2	135 ± 12	19 ± 4	4 ± 2
Single top	4953 ± 70	339 ± 18	8 ± 3	1849 ± 43	191 ± 14	9 ± 3	734 ± 27	115 ± 11	11 ± 3
Diboson	167 ± 13	9 ± 3	0 ± 0	80 ± 9	7 ± 3	0 ± 1	39 ± 6	5 ± 2	0 ± 1
Z + jets	474 ± 22	25 ± 5	0 ± 1	175 ± 13	15 ± 4	1 ± 1	54 ± 7	5 ± 2	0 ± 0
W + jets	2693 ± 52	126 ± 11	3 ± 2	1062 ± 33	76 ± 9	4 ± 2	507 ± 23	59 ± 8	5 ± 2
$t\bar{t} + V$	98 ± 10	13 ± 4	1 ± 1	113 ± 11	22 ± 5	3 ± 2	136 ± 12	35 ± 6	7 ± 3
$t\bar{t} + b\bar{b}$	971 ± 31	442 ± 21	37 ± 6	982 ± 31	593 ± 24	96 ± 10	942 ± 31	789 ± 28	214 ± 15
$t\bar{t} + c\bar{c}$	3405 ± 58	626 ± 25	25 ± 5	3253 ± 57	748 ± 27	45 ± 7	2895 ± 54	769 ± 28	81 ± 9
$t\bar{t}$ + light	66 042 ± 257	5468 ± 74	47 ± 7	32 677 ± 181	3255 ± 57	59 ± 8	16 224 ± 127	1958 ± 44	58 ± 8
Total background	79 677 ± 282	7112 ± 84	124 ± 11	40 462 ± 202	4948 ± 70	221 ± 15	21 665 ± 147	3755 ± 61	380 ± 19
$t\bar{t}H$ signal	100 ± 10	42 ± 7	7 ± 3	128 ± 11	71 ± 8	19 ± 4	191 ± 14	126 ± 11	51 ± 7
Data	79 598	7140	128	40 849	5088	232	21 705	3901	433

Table 7.7: Table showing the yields of each process in each analysis region after the fit to data, rounded to the nearest whole event.

Region	(4,2)	(4,3)	(4,4)	(5,2)	(5,3)	(5,4)	(6,2)	(6,3)	(6,4)
Multijet	0.42	0.42	0.43	0.42	0.42	0.45	0.42	0.41	0.44
Single top	0.97	1.01	1.14	0.96	0.99	1.00	0.94	1.02	1.00
Diboson	0.99	1.00	0.00	1.01	1.00	0.00	0.93	1.00	0.00
$Z+$jets	0.98	1.04	0.00	0.96	1.00	1.00	1.00	1.00	0.00
$W+$jets	0.97	1.01	1.00	0.97	1.01	1.33	0.96	1.02	1.00
$t\bar{t}+V$	1.00	1.00	1.00	0.98	1.00	1.00	0.96	0.95	1.00
$t\bar{t}+b\bar{b}$	1.07	1.08	1.12	1.04	1.06	1.09	0.99	1.08	1.00
$t\bar{t}+c\bar{c}$	1.21	1.26	1.32	1.19	1.24	1.12	1.14	1.14	1.23
$t\bar{t}+$light	1.02	1.05	1.02	1.03	1.06	1.05	1.00	1.06	1.04
Total background	1.01	1.05	1.06	1.02	1.06	1.05	1.01	1.07	1.03
$t\bar{t}H$ signal	3.70	3.82	3.50	3.66	3.74	3.80	3.54	3.71	3.64

Table 7.8: Table showing the pre-to-post-fit ratio of the median yield of each process in each region (as in Tables 7.6 and 7.7), rounded to two decimal places.

95 % CL upper limit value	Statistical error only	Systematic & statistical
$+2\sigma$	3.4	8.3
$+1\sigma$	2.5	5.9
median	1.8	4.2
-1σ	1.3	3.0
-2σ	1.0	2.2
observed	—	5.9

Table 7.9: Table showing the expected limits on Standard Model $t\bar{t}H, H \to b\bar{b}$ for $m_H = 125\,\text{GeV}$. All values represent $\sigma/\sigma_{\text{SM}}$.

Table 7.9 shows the expected limit on $t\bar{t}H, H \to b\bar{b}$ for the mass-point $m_H = 125\,\text{GeV}$, together with the $\pm 1, 2\sigma$ uncertainties, expressed as a ratio of the expected or observed cross-section to the Standard Model cross-section. The expected median limit was found to be $4.2^{+1.7}_{-1.2}$, where the statistical and systematic uncertainties are combined, and the observed limit was 5.9, with an observed signal strength of $\mu = 2.05 \pm 2.78$. It is notable that, likely due to the slightly stricter event selection in this analysis and the lesser signal-background separation power of the neural networks trained in TMVA within this analysis as compared with NeuroBayes, these limits are slightly higher than – but consistent with – those of the ATLAS $t\bar{t}H, H \to b\bar{b}$ analysis, which observed a limit of $\sigma/\sigma_{\text{SM}} = 4.1$ [93].

Chapter 8

Search for a Heavy Charged Higgs Boson

"Hypotheses non fingo."

— Newton, *Philosophiæ Naturalis Principia Mathematica*

8.1 Search for a Heavy Charged Higgs

The analysis techniques outlined in Chapter 7 were extended to search for a heavy charged Higgs $H^{\pm}$ of mass $m_{H^{\pm}} = 200, 250, 300, 400, 500, 600\,\text{GeV}$, working within the five-flavour scheme with the decay mode $H^{\pm} \to tb$. This is a beyond-the-Standard-Model (BSM) process with a similar final state to that of $t\bar{t}H, H \to b\bar{b}$, as illustrated in Figure 2.6 (page 23), with a different signal contribution to the expected b-jet multiplicity (three b-jets in this BSM scenario, as opposed to four in the Standard Model case). There is thus a reasonable expectation that the analysis techniques which have been employed in the search for the Standard Model $t\bar{t}H$ signal will be relevant in searching for $H^{\pm} \to tb$.

For this analysis, the analysis regions defined for the Standard Model $t\bar{t}H, H \to b\bar{b}$ search were adapted so that the same total event-space was partitioned into five regions: the control (signal-depleted) regions (4,2), (5,2), and (6,2) remained unchanged from the Standard Model analysis, while the regions (4,3) and (4,4) were combined to form a single control region, hereafter referred to as (4,3) (denoting 4 jets and *at least* 3 b-tags) in the context of the charged-Higgs analysis. There was a single signal-enriched region, denoted (5,3), here representing *at least* 5 jets and *at least* 3 b-tags. This signal-enriched region is therefore equivalent to the union of the

four Standard Model regions (5,3), (5,4), (6,3), and (6,4). These analysis regions are illustrated in Table 8.1. As in the Standard Model analysis, a multilayer perceptron neural network was trained using TMVA within this signal-enriched region in order to optimise separation between the $H^{\pm} \to tb$ signal and the $t\bar{t}b\bar{b}$ background, which also serves as the main contribution to the background in this analysis. The input variables to TMVA training are shown in Table 8.2, and the signal-background separations attained by the TMVA following training at each mass point are shown in Figure 8.1.

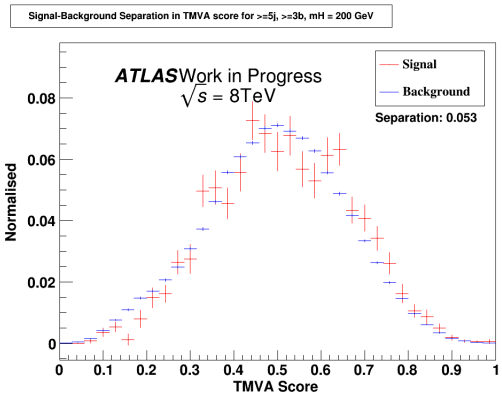
	n_j		
n_b	4	5	≥ 6
2	BR	BR	BR
3	BR	SR	SR
≥ 4	BR	SR	SR

Table 8.1: Table outlining the analysis regions used in the charged-Higgs analysis. "BR" denotes background (signal-depleted) regions while "SR" denotes signal-enriched regions. In contrast to the Standard Model analysis region definitions as shown in Table 7.1, the four regions denoted "SR" here constitute a single combined signal-enriched region.

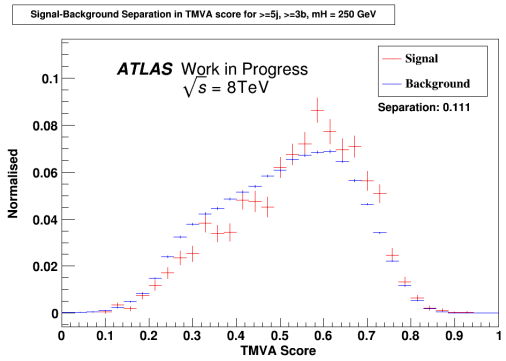
The data sample and the background Monte Carlo samples are as described in Chapter 7. The charged-Higgs signal samples are prepared using POWHEG with the CT10 NLO PDF set, interfaced to PYTHIA 8. The charged-Higgs decays were simulated in PYTHIA 8, with a one-lepton filter (without a truth-level p_{T} cut) on the decay products of the W boson so as to select lepton-plus-jets events only. Since the expected width of the charged Higgs boson is far below experimental resolution, it is set to zero.

Variable	Definition
H_1	Fox-Wolfram moment of order 1
$H_{\mathrm{T}}^{\mathrm{had}}$	Scalar sum of all jet p_{T}
m_{jjj}	Mass of the jet triplet with the largest vector-sum p_{T}
$m_{W_{\mathrm{had}}}$	Reconstructed mass of the hadronic W boson
aplanarity$_{b-\mathrm{jet}}$	Equal to $1.5\lambda_2$ (λ_2 is the second eigenvalue of the momentum tensor constructed using only b-jets)
centrality	centrality $= \Sigma p_{\mathrm{T}}/\Sigma E$ (summing over all jets and the lepton)
$m_{\mathrm{bj}}^{\mathrm{max}p_{\mathrm{T}}}$	Invariant mass of the pair comprising a b-jet and any jet, chosen so as to maximise pair p_{T}
N_{40}^{jet}	Number of jets in event satisfying $p_{\mathrm{T}} > 40\,\mathrm{GeV}$
$p_{\mathrm{T}}^{\mathrm{jet5}}$	p_{T} of the fifth-leading jet
$p_{\mathrm{T}}^{jj,\mathrm{av}}$	Average p_{T} of a jet-jet pair

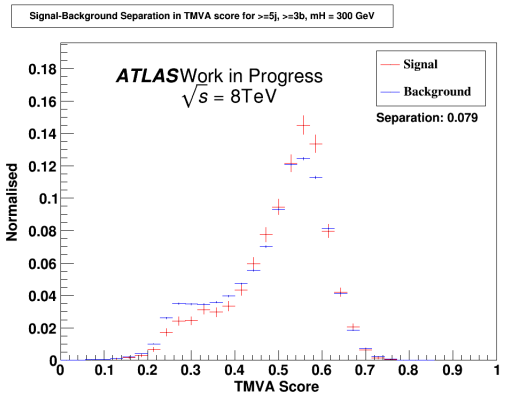
Table 8.2: Table listing input variables to TMVA training for the charged-Higgs analysis.



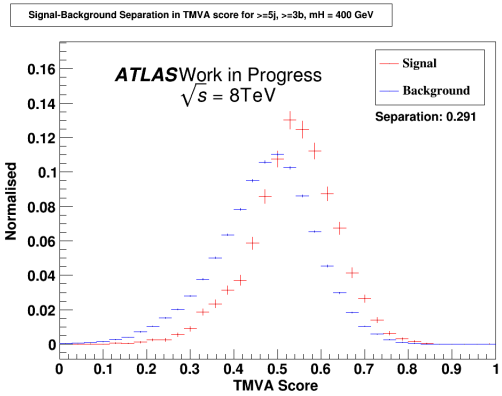
(a) $m_{H^\pm} = 200\,\text{GeV}$



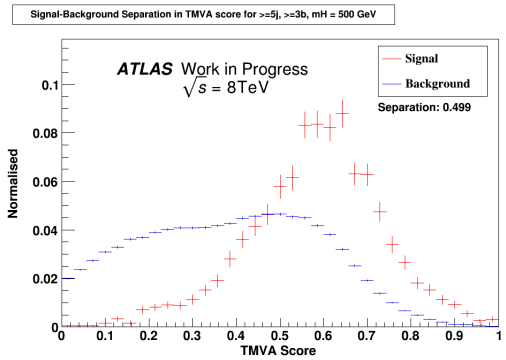
(b) $m_{H^\pm} = 250\,\text{GeV}$



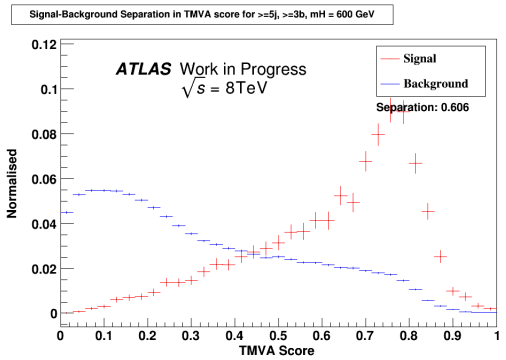
(c) $m_{H^\pm} = 300\,\text{GeV}$



(d) $m_{H^\pm} = 400\,\text{GeV}$



(e) $m_{H^\pm} = 500\,\text{GeV}$



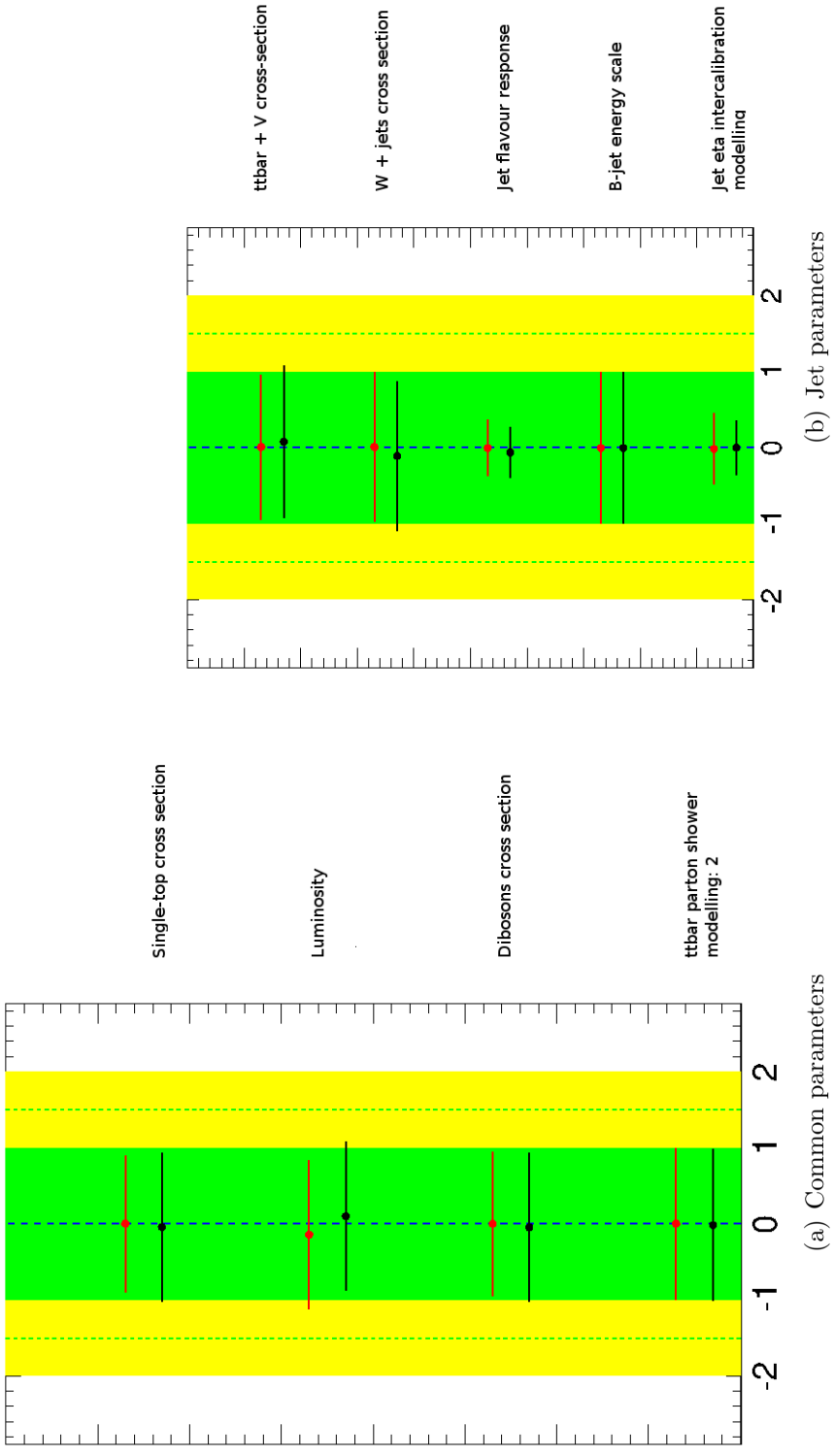
(f) $m_{H^\pm} = 600\,\text{GeV}$

Figure 8.1: Plots showing the shape difference and signal-background separation in the (5,3) region for each charged-Higgs mass point.

The treatment of the systematic uncertainties was simplified by considering only the highest-ranked contributions to the total systematic uncertainty, according to the ATLAS $H^{\pm} \to tb$ analysis [179], which are listed in Table 8.3. Following the prescription of the ATLAS group, an uncertainty is not applied to the signal Monte Carlo rate, as the effects on the limit of scale variation, generator variation and PDF scale were found by them to be negligible for all mass points. Moreover, it was found by the ATLAS group that test samples produced in MadGraph using the four-flavour scheme gave consistently higher signal yields than those using the five-flavour scheme, and therefore the limits obtained using the 5FS signal Monte Carlo are to be taken as conservative.

Systematic Name	Type
Luminosity	N
$t\bar{t}$ cross section	N
$t\bar{t}b\bar{b}$ cross section	N
$t\bar{t}c\bar{c}$ cross section	N
W+jets cross section	N
Z+jets cross section	N
$t\bar{t}+V$ cross section	N
Multijet cross section	N
Dibosons cross section	N
Single-top cross section	N
b-jet tagging: component 4	SN
b-jet tagging: component 5	SN
c-jet tagging: component 1	SN
c-jet tagging: component 2	SN
c-jet tagging: component 3	SN
c-jet tagging: component 4	SN
c-jet tagging: component 5	SN
c-jet tagging: component 6	SN
JES: Jet η intercalibration	SN
b-jet energy scale	SN
Mixture of light-quark and gluon jets	SN
Light jet tagging: component 11	SN

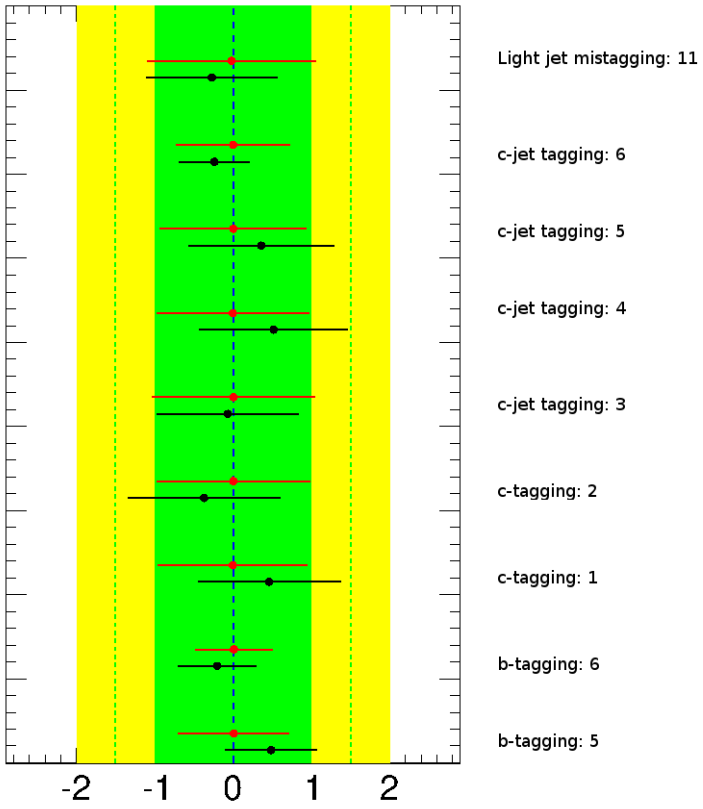
Table 8.3: Table showing the sources of systematic uncertainty considered in the charged-Higgs analysis; these form a subset of those used in the Standard Model analysis, which are detailed in Table 7.2. A number following the name of a systematic denotes one of the independent components of a given source of systematic uncertainty (following the ranking convention adopted by the ATLAS $t\bar{t}H$ group), and "S" and "N" denote whether a given source of systematic uncertainty is considered as affecting the shape or normalisation, or both, of the Monte Carlo.



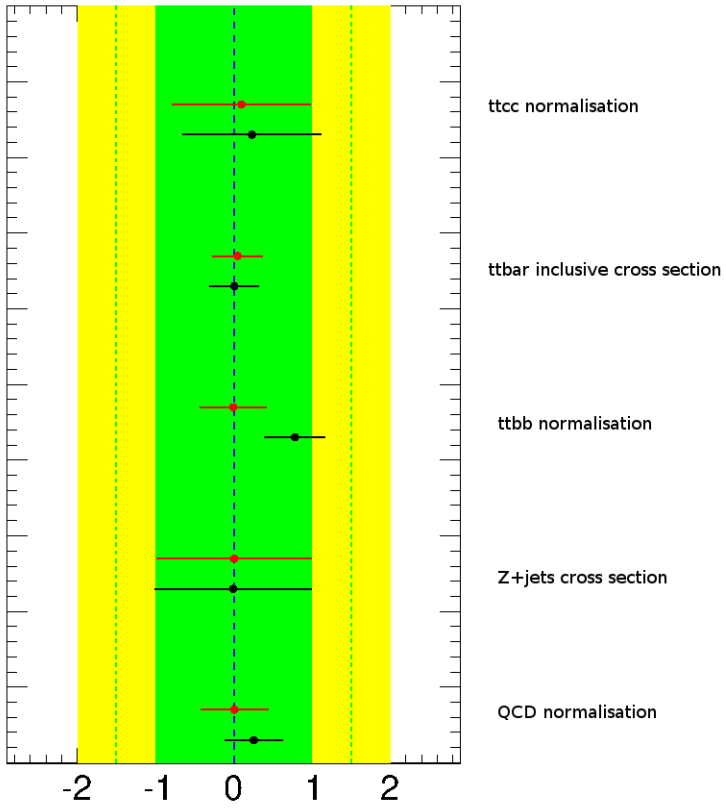
(a) Common parameters

(b) Jet parameters

Figure 8.2: Nuisance parameters fitted to Asimov data (red) and data (black) across all analysis regions for the signal-plus-background hypothesis, for the mass-point $m_{H^\pm} = 400\,\text{GeV}$

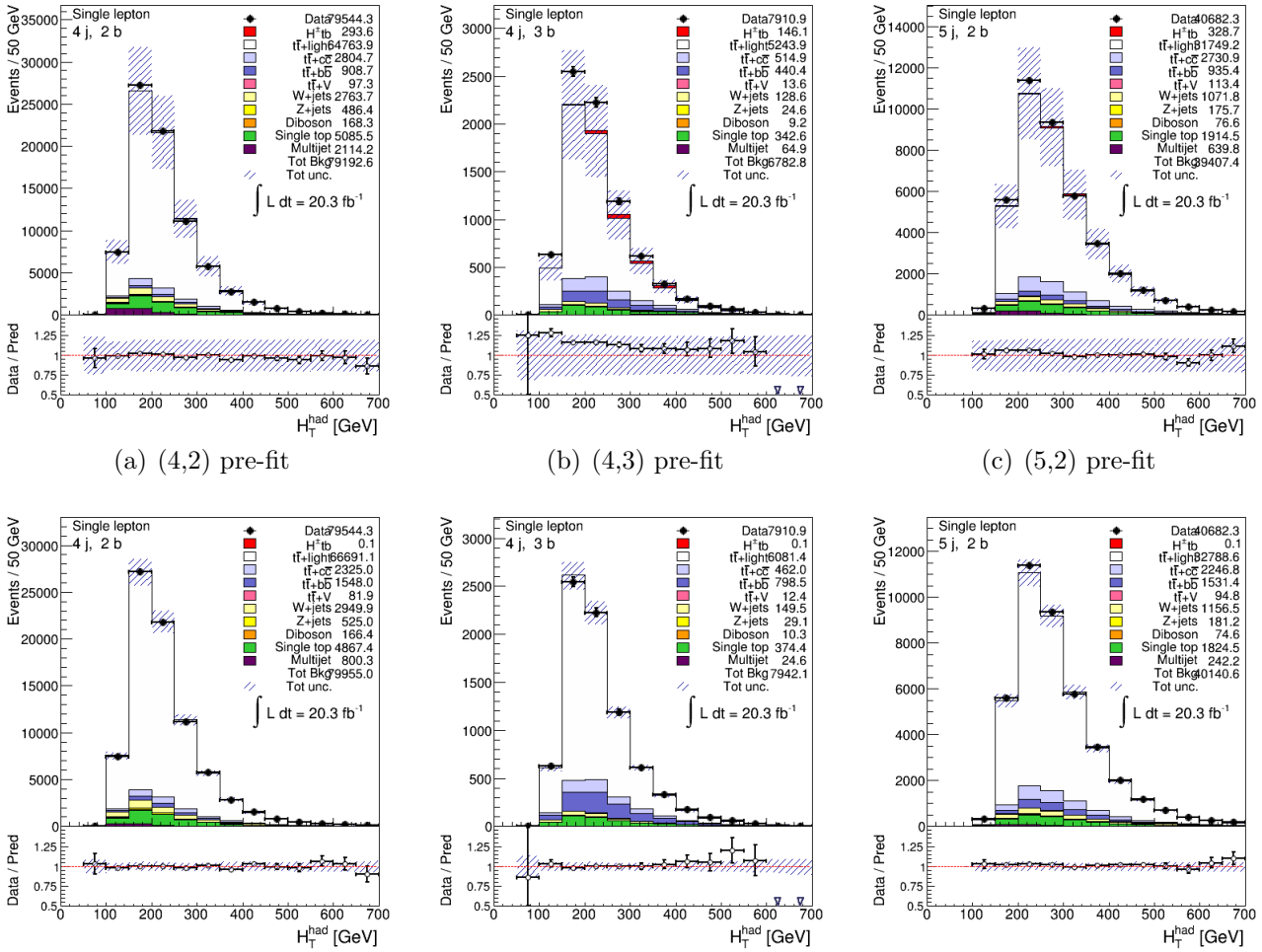


(a) Lepton and b parameters



(b) Other parameters

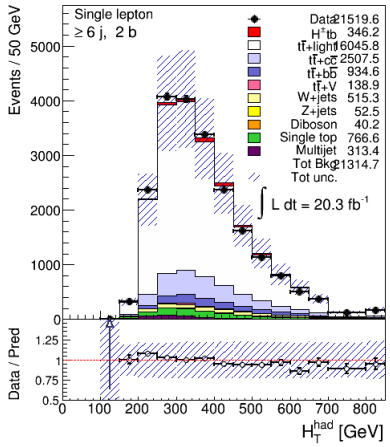
Figure 8.3: Nuisance parameters fitted to Asimov data (red) and data (black) across all analysis regions for the signal-plus-background hypothesis, for the mass-point $m_{H^\pm} = 400\,\text{GeV}$



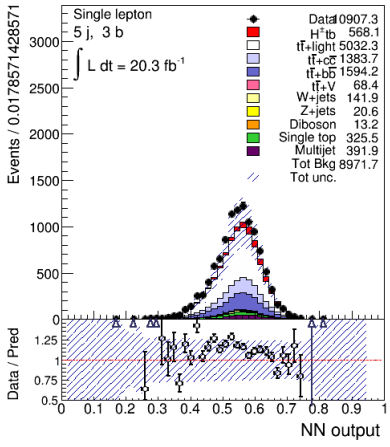
(a) (4,2) pre-fit (b) (4,3) pre-fit (c) (5,2) pre-fit

(d) (4,2) post-fit (e) (4,3) post-fit (f) (5,2) post-fit

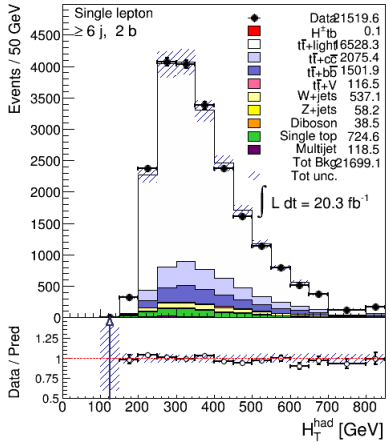
Figure 8.4: Plots showing the stacked contributions to each bin, before and after the fit to data, in the regions (4,2), (4,3), and (5,2), for the mass-point $m_{H^\pm} = 400\,\text{GeV}$.



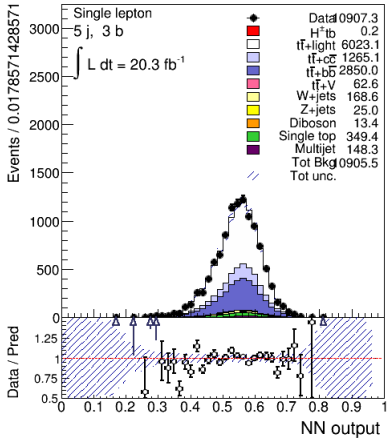
(a) (6,2) pre-fit



(b) (5,3) pre-fit



(c) (6,2) post-fit



(d) (5,3) post-fit

Figure 8.5: Plots showing the stacked contributions to each bin, before and after the fit to data, in the regions (5,3) and (6,2), for the mass-point $m_{H^\pm} = 400\,\text{GeV}$.

For each mass-point under consideration, the GlaNtp analysis framework was used to produce fit templates for each systematic uncertainty in each analysis region; these templates then served as inputs to the limit-setting code. The expected and observed limits were thus computed by the same procedure as was used in the Standard Model analysis presented in Chapter 7 – these limits are displayed in Figure 8.6. A small excess of data over expectation is observed for $m_{H^\pm} = 200, 250, 300\,\text{GeV}$, with a peak at 250 GeV. For the higher charged-Higgs masses $m_{H^\pm} = 400, 500, 600\,\text{GeV}$, the fit to data produces results which are slightly less compatible with the signal-plus-background hypothesis when compared to the corresponding expected limits.

The local p-values corresponding to the statistical significance S of the observed limits on these charged-Higgs scenarios are shown in Figure 8.7; these are calculated according to Equation (6.8). It can be seen that the lowest p-value (corresponding to the observation with the lowest conditional probability given the signal-plus-background hypothesis at its mass point), is obtained for $m_{H^\pm} = 250\,\text{GeV}$. Conversely, the mass point for which the observed limit is closest to the expected limit, $m_{H^\pm} = 500\,\text{GeV}$, yields the highest p-value.

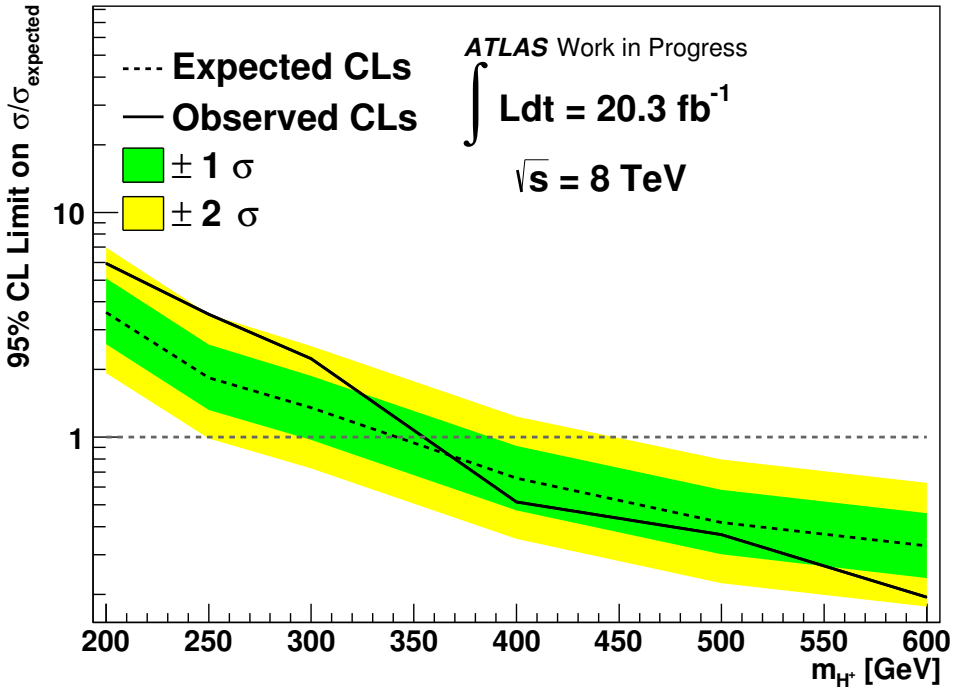


Figure 8.6: The expected and observed 95% CL limits on $H^\pm$ production for the mass range $200\,\text{GeV} \leq m_{H^\pm} \leq 600\,\text{GeV}$. σ_{expected} denotes the expected $\sigma \times \mathcal{B}$ as computed for each value of $m_{H^\pm}$.

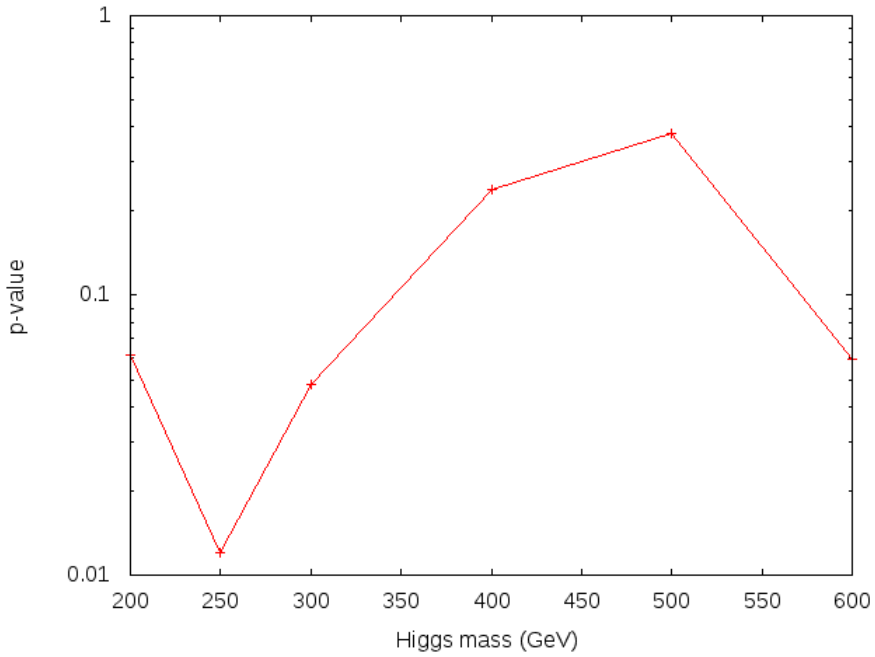


Figure 8.7: The p-values corresponding to the significance of the observed limits in Figure 8.6.

8.2 Conclusions

A fit of Monte Carlo to ATLAS $\sqrt{s} = 8\,\text{TeV}$ data has been performed in order to compute limits on the charged Higgs decay mode $H^{\pm} \to tb$, using the charged-Higgs mass points $m_{H^{\pm}} = 200, 250, 300, 400, 500, 600\,\text{GeV}$. A very slight excess of events above the background-only expectation was observed for the mass point $m_{H^{\pm}} = 250\,\text{GeV}$; for all other mass points, the observed limit was statistically compatible (to within the 95 % confidence level) with the background-only hypothesis. For mass points $m_{H^{\pm}} \geq 400\,\text{GeV}$, the observed limits fell somewhat below the expected limits. These findings are broadly consistent with those of the ATLAS analysis presented in Ref. [179].

Chapter 9

Summary and Conclusions

> "I was one, sir, in this interlude, one Sir Topas, sir, but that's all one."
>
> — Shakespeare, *Twelfth Night*

By the end of its first run in February 2013, the LHC had accomplished one of the major goals of experimental physics in the four decades that preceded it – the discovery of the Higgs boson, with which the Standard Model "zoo" gained its most elusive member.

2015 has seen the beginning of the second run of the LHC, at a greatly increased centre-of-momentum collision energy of 13 TeV, and with higher instantaneous luminosity due to the decrease in bunch spacing to the LHC's design capacity of 25 ns. This run has the prospect to discover yet more new fundamental physics.

In the current era of scientific collaborations with thousands of members and ever more complex organisational structures, when competitiveness rewards swift completion and publication of analyses, it is more important than ever before to provide for the performance of independent cross-checks of experimental results. The work presented in this thesis represents one individual effort to carry out an analysis parallel to an official ATLAS analysis, using an alternative analysis software framework and an independent workflow.

By careful and consistent choices of parameters and settings, the MadGraph Monte Carlo generator was successfully validated against Alpgen for the production of samples of the background process $t\bar{t}b\bar{b}$ suitable for use in the Standard Model $t\bar{t}H, H \to b\bar{b}$ analysis.

In the search for the Standard Model process $t\bar{t}H, H \to b\bar{b}$, it was found that the

signal-background discrimination power achievable with the chosen method – TMVA multilayer perceptron training implemented within the GlaNtp analysis framework – was not as great as that achieved by NeuroBayes within the published ATLAS analysis [93], and therefore the observed limit of $\sigma/\sigma_{\mathrm{SM}} = 5.9$ was correspondingly higher than that of $\sigma/\sigma_{\mathrm{SM}} = 4.2$ as set by the ATLAS group. A natural extension of this work would be to compare the separation power achievable by other multivariate techniques (such as a boosted decision tree), and the corresponding effects of these techniques on the observed limit.

A simplified extension of the Standard Model $t\bar{t}H, H \to b\bar{b}$ analysis to set limits on the process $H^{\pm} \to tb$ produced results comparable to those of the published ATLAS analysis [179], with a slightly smaller observed excess overall, and an observed limit close to, but slightly below, the expected limit for mass points $m_{H^{\pm}} \geq 400\,\mathrm{GeV}$. Further limit-setting studies on this process could also benefit from considering alternative multivariate techniques as means of computing discriminating variables.

Appendices

Appendix A

MadGraph Scale Validation

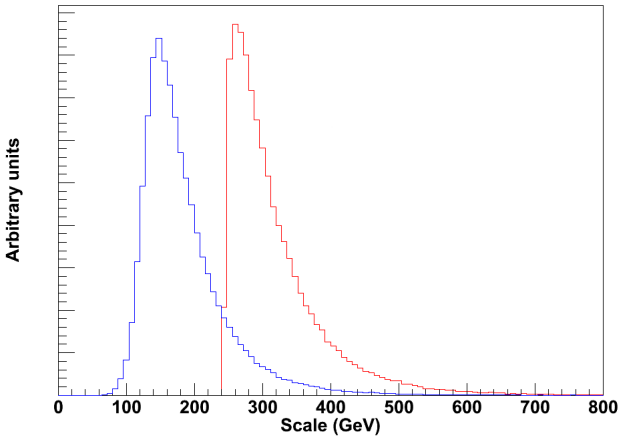
> "There are things we know that we know. There are known unknowns. That is to say there are things that we now know we don't know. But there are also unknown unknowns. There are things we do not know we don't know."
>
> — Donald Rumsfeld

It was desirable to validate the implementation of the Alpgen scale in the generation of the MadGraph samples at the parton level, *i.e.* prior to showering and the Rivet analysis. It was anticipated that a choice of scale should not in principle lead to significant differences in distributions of observables, as the scale is in essence a theoretical construct; all that was important in the studies described in Chapter 5 was that a single scale was implemented consistently in order to ensure a fair comparison.

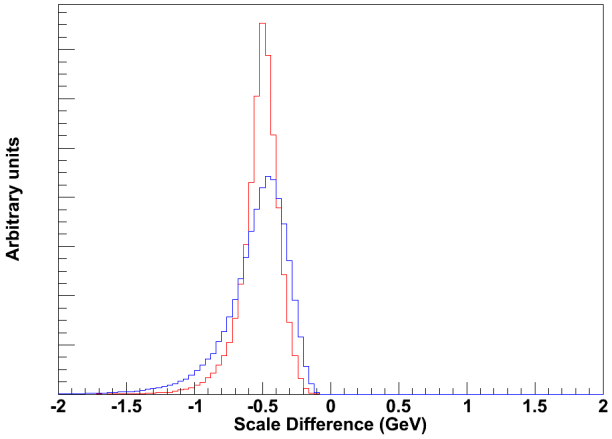
Figure A.1(a) shows the effect of disabling top decays on the scale value as recorded in the MadGraph LHEF output – a large discrepancy is observed. Since the MadGraph scale is defined with respect to the maximum mass of a final-state particle, this discrepancy was to be expected; disabling the top decays would imply that the most massive final-state particle would be the top itself (which is far more massive than its daughter products).

Figures A.1(b) and A.1(c) detail a validation exercise to ensure that the Alpgen scale was correctly implemented in MadGraph. For each event in the LHEF sample, the scale was computed indepedently of MadGraph using the Alpgen scale formula, as shown in equation (5.7); the difference between this computed quantity and the nominal scale value was then taken. Figure A.1(b) shows the absolute dif-

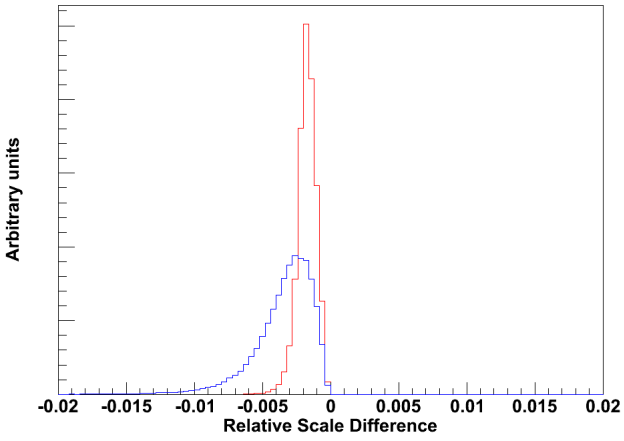
ference between these two values of the scale, while Figure A.1(c) shows the relative difference (*i.e.* the absolute difference divided by the value of the scale as computed by MadGraph). It is notable that, while extremely low (around 0.5 GeV), the difference between the recorded and calculated scales does not peak at zero.



(a) Comparison of Alpgen and MadGraph scales



(b) Absolute difference in scale



(c) Relative difference in scale

Figure A.1: Comparison of scales in MadGraph LHEF files, with and without top decays. The red histograms contain $t\bar{t}b\bar{b}$ events without top decays at the generator level, while the blue histograms contain otherwise identical events with top decays enabled in MadGraph and without any cuts being imposed on the top decay products.

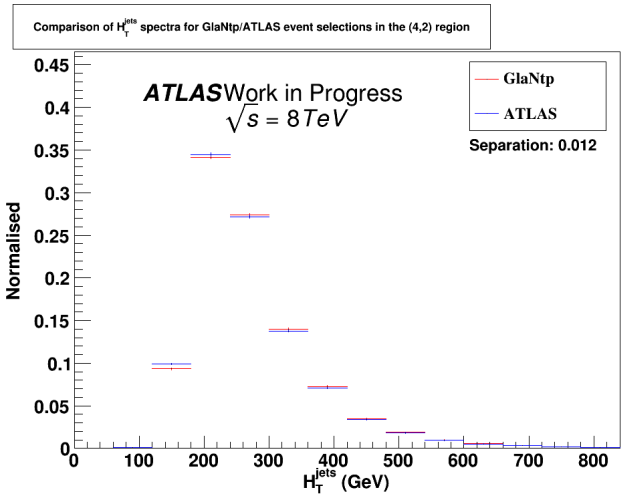
Appendix B

GlaNtp Analysis Cuts in the Standard Model $t\bar{t}H$ Analysis

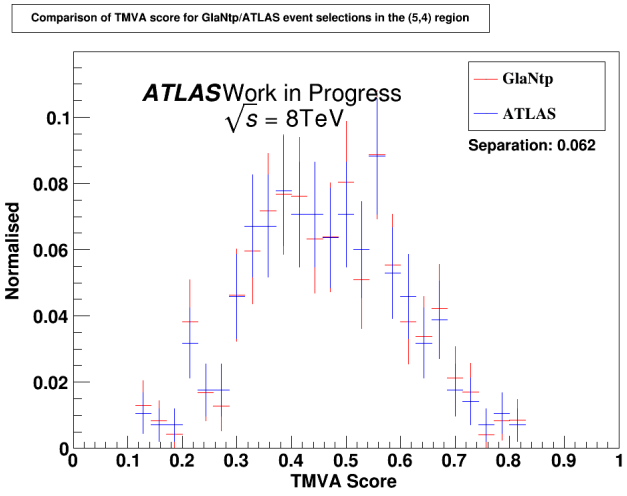
"E sarà mia colpa, se così è?"

— Stendhal, *Le Rouge et le Noir*

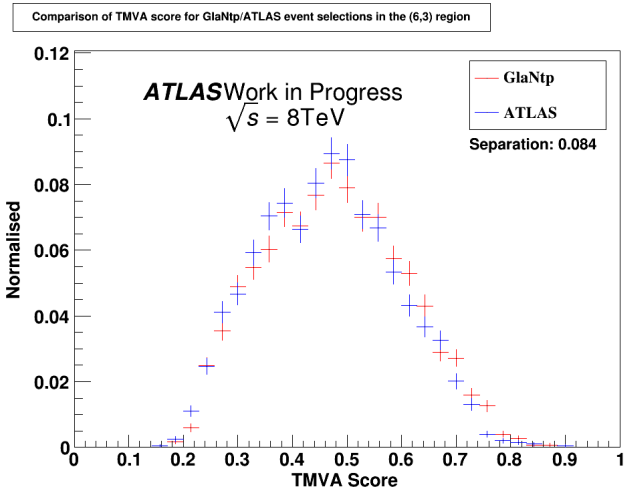
The event selection implemented in the GlaNtp analysis package as described in Chapter 7 differed somewhat from that used in the ATLAS $t\bar{t}H$ working group's analysis as documented in Ref. [93] – in GlaNtp an additional cut of $E_{\mathrm{T}}^{\mathrm{miss}} \geq 25$ GeV was applied, which reduced the observed yields across the analysis by approximately 15–20 % for Monte Carlo and data with respect to the published ATLAS analysis. As a verification exercise, in order to ensure that significant shape differences were not introduced into the spectra of key observables, the effects of this additional cut on the shapes of the $H_{\mathrm{T}}^{\mathrm{jets}}$ spectrum in the (4,2) analysis region and on the TMVA score distribution in the those analysis regions where a neural network is trained were examined; these are shown in Figure B.1. As can be seen, large shape differences in the fit variables are not induced by the application of this cut, suggesting that the GlaNtp selection is consistent with (although stricter than) that of the published ATLAS analysis.



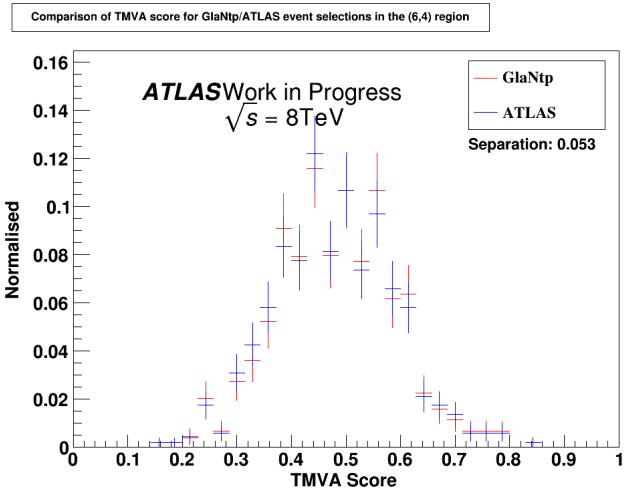
(a) $H_{\mathrm{T}}^{\mathrm{jets}}$, (4,2)



(b) TMVA score, (5,4)



(c) TMVA score, (6,3)



(d) TMVA score, (6,4)

Figure B.1: Data-data comparison plots showing the difference in the shape of the $H_{\mathrm{T}}^{\mathrm{jets}}$ spectrum in the control region (4,2) and the TMVA score distributions in the signal-enriched regions (5,4), (6,3), and (6,4).

Appendix C

GlaNtp Input Variable Plots

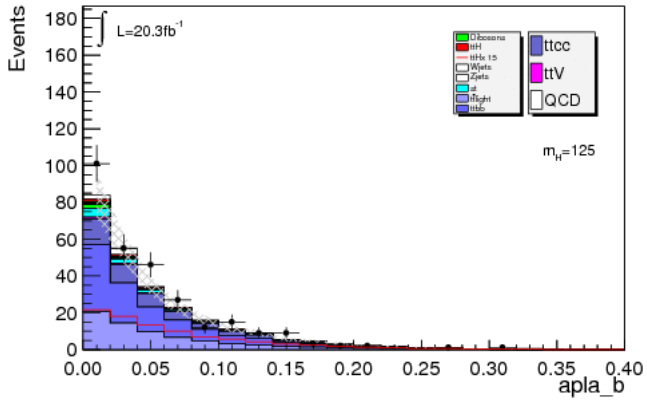
"I try all things; I achieve what I can."

— Herman Melville, *Moby-Dick; or, The Whale*

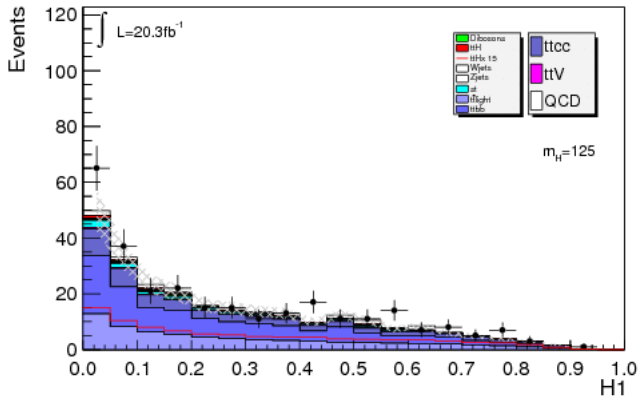
GlaNtp includes functionality to allow it to plot the input variables to neural network training. Stacked plots of the three highest-ranked (by TMVA importance) variables in the signal-enriched region (5,4) are shown in Figure C.1; those for the region (6,3) are shown in Figure C.2, and those for the region (6,4) are shown in Figure C.3. These rankings are as shown in Table 7.4. For the sake of clarity, the processes shown in the smaller legend box in each of the following figures are described in Table C.1.

Colour	Process
Green	Dibosons
Red (solid)	$t\bar{t}H$
Red (line)	$t\bar{t}H \times 15$
White	W+jets
Beige	Z+jets
Aquamarine	Single top
Light violet	$t\bar{t}$+light jets
Dark violet	$t\bar{t}b\bar{b}$

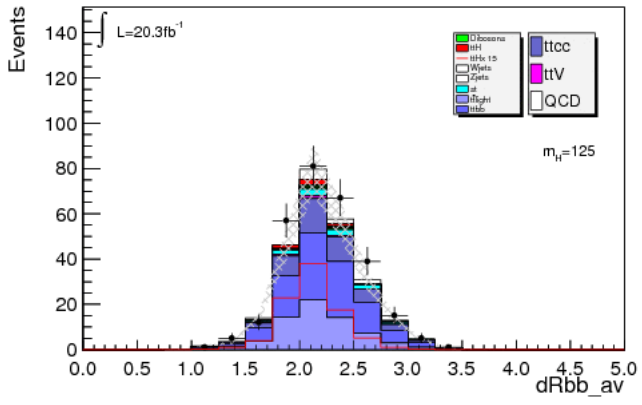
Table C.1: The colour scheme used to denote signal and background processes in the figures in this appendix. These are the processes shown in the smaller legend box.



(a) (5,4): aplanarity$_{b-\text{jet}}$

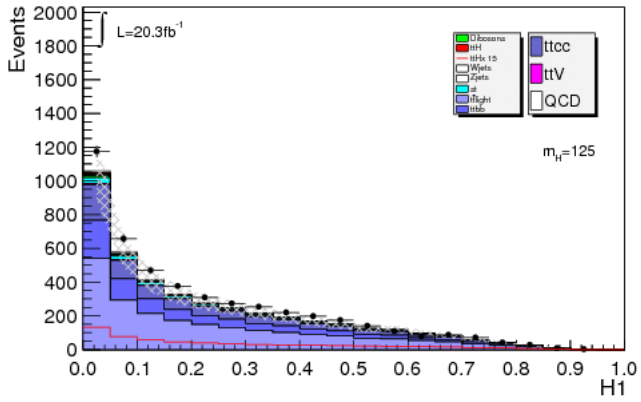


(b) (5,4): H_1

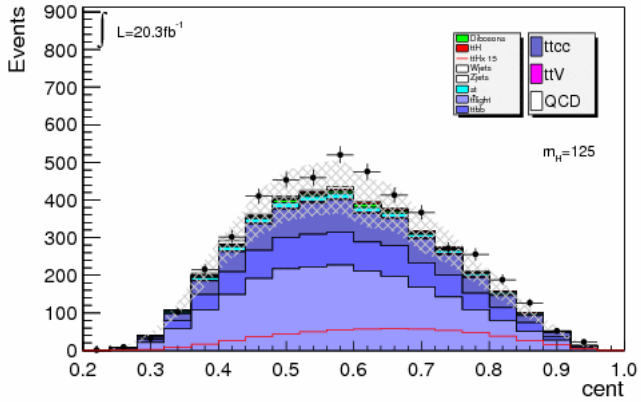


(c) (5,4): $\Delta R_{bb}^{\text{avg}}$

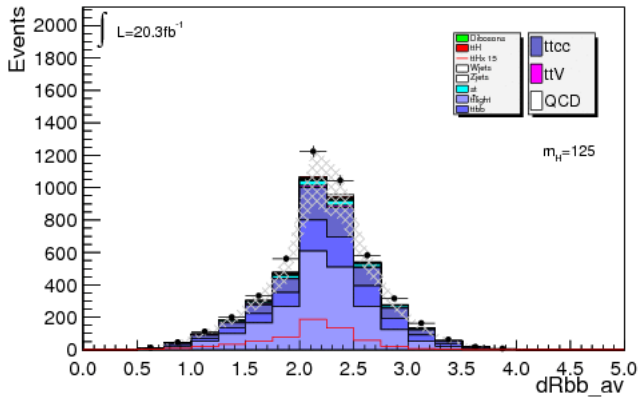
Figure C.1: Plots showing the distributions of the leading three variables (as ranked by TMVA importance) to TMVA training in the Standard Model analysis region (5,4). The $t\bar{t}H$ signal is scaled up by a factor of 15 for ease of visualisation.



(a) (6,3): H_1

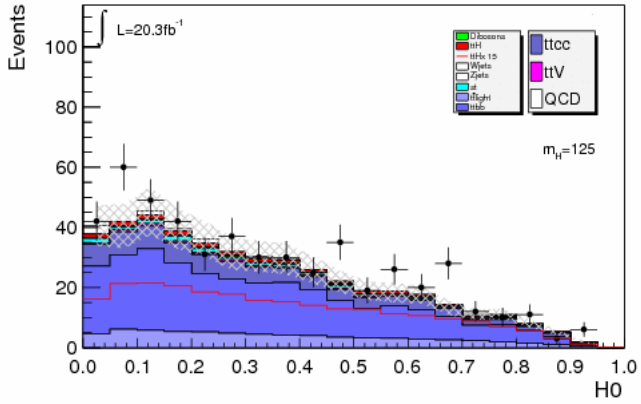


(b) (6,3): centrality

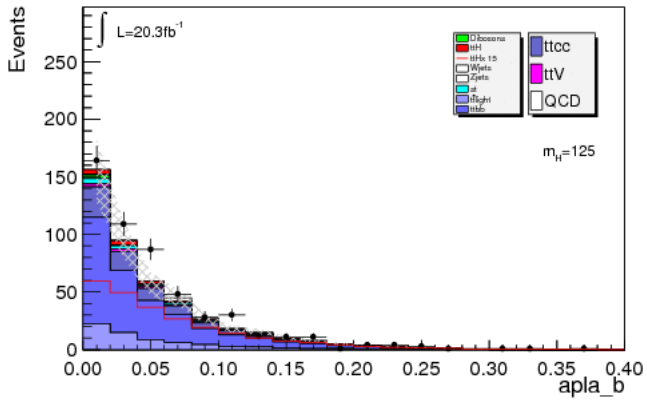


(c) (6,3): $\Delta R_{bb}^{\text{avg}}$

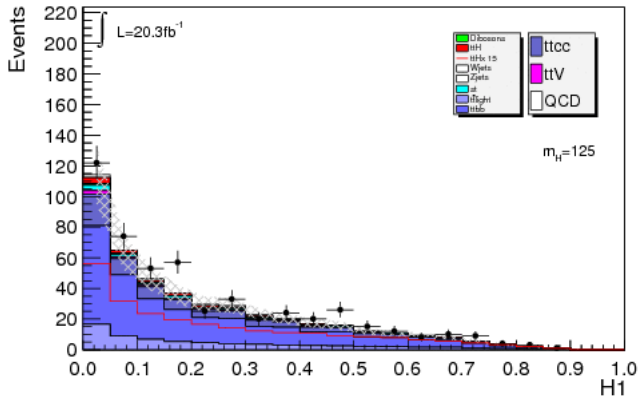
Figure C.2: Plots showing the distributions of the leading three variables (as ranked by TMVA importance) to TMVA training in the Standard Model analysis region (6,3). The $t\bar{t}H$ signal is scaled up by a factor of 15 for ease of visualisation.



(a) (6,4): H_0



(b) (6,4): aplanarity$_{b-\text{jet}}$



(c) (6,4): H_1

Figure C.3: Plots showing the distributions of the leading three variables (as ranked by TMVA importance) to TMVA training in the Standard Model analysis region (6,4). The $t\bar{t}H$ signal is scaled up by a factor of 15 for ease of visualisation.

Appendix D

Glossary for SM Analysis Nuisance Parameter Plots

"I understand a fury in your words, But not the words."

— Shakespeare, *Othello*

This appendix is a key to the abbreviated names of the nuisance parameters used in the pull distributions shown in Figures 7.6–7.9. These are detailed in Table D.1. The 22 components of jet energy scale (JES) are as provided to the ATLAS $t\bar{t}H, H \to b\bar{b}$ analysis group by the ATLAS JES working group. Where component numbering is used in the table, it reflects the numbering conventions of the ATLAS JES group.

Abbreviation	Meaning
ttbar reweighting: 1	$t\bar{t}$ data reweighting – jet η calibration
ttbar reweighting: 2	$t\bar{t}$ data reweighting – jet detector effects
ttbar reweighting: 3	$t\bar{t}$ data reweighting – jet overlap
ttbar reweighting: 4	$t\bar{t}$ data reweighting – b-tag efficiency
ttbar reweighting: 5	$t\bar{t}$ data reweighting – b-jet energy scale
ttbar reweighting: 6	$t\bar{t}$ data reweighting – MC generator
ttbar reweighting: 7	$t\bar{t}$ data reweighting – jet energy resolution
ttbar reweighting: 8	$t\bar{t}$ data reweighting – initial- & final-state radiation
ttbar reweighting: 9	$t\bar{t}$ data reweighting – fragmentation
ttbar + HF: 1	Generator choice (PowHeg versus MadGraph)
ttbar + HF: 2	Energy scale (Q^2) variation in MadGraph
ttbar + HF: 3	c-mass variation in MadGraph
ttbar + HF: 4	b-mass variation in MadGraph
ttbar + HF: 5	Heavy-flavour matching in MadGraph
ttbar + HF: 6	Parton shower modelling: light flavours

ttbar + HF: 7	Parton shower modelling: heavy flavours
JES: 1	Effective JES statistical/method component 3
JES: 2	Effective JES statistical/method component 2
JES: 3	Effective JES statistical/method component 1
JES: 4	Effective JES modelling component 4
JES: 5	Effective JES modelling component 3
JES: 6	Effective JES modelling component 2
JES: 7	Effective JES modelling component 1
JES: 8	Effective JES mixed modelling/detector component 2
JES: 9	Effective JES mixed modelling/detector component 1
JES: 10	Effective JES detector component 3
JES: 11	Effective JES detector component 2
JES: 12	Effective JES detector component 1
JES: 13	Mixture of light-quark and gluon jets
JES: 14	JES uncertainty for gluon jets
JES: 15	Abundance of single-parton jets
JES: 16	Pileup modelling – jet p_{T}
JES: 17	Density of pileup contamination (ρ)
JES: 18	Pileup – offset, no. of primary vertices ($\mathrm{N_{PV}}$) term
JES: 19	Pileup – offset, pileup jet multiplicity (μ) term
JES: 20	Calorimeter response to b-jets
JES: 21	Jet eta intercalibration – data statistics
JES: 22	Jet eta intercalibration – modelling
Electrons: 1	Electron identification
Electrons: 2	Electron resolution
Electrons: 3	Electron trigger
Electrons: 4	Electron energy scale
Electrons: 5	Electron reconstruction
Muons: 1	Muon trigger
Muons: 2	Muon reconstruction
Muons: 3	Muon identification
Muons: 4	Muon momentum scale
Muons: 5	Muon resolution in the muon system
Muons: 6	Muon resolution in the tracker
ttH modelling: 1	Functional form of $t\bar{t}H$ scale
ttH modelling: 2	Scale variation
ttbar parton shower modelling: 1	Light flavours
ttbar parton shower modelling: 2	Heavy flavours

Table D.1: A glossary of the abbreviated terms used in the nuisance parameter plots for the Standard Model $t\bar{t}H, H \to b\bar{b}$ analysis.

Bibliography

[1] Mike Lamont and Cian O'Luanaigh. *LHC report: Run 1 – the final flurry*. Feb. 2013. URL: http://cds.cern.ch/record/1997755 (cit. on p. 1).

[2] Amos Breskin and Rüdiger Voss. *The CERN Large Hadron Collider: Accelerator and Experiments*. Geneva: CERN, 2009. URL: https://cds.cern.ch/record/1244506 (cit. on pp. 1, 29, 31).

[3] Oliver Sim Brüning et al. *LHC Design Report*. Geneva: CERN, 2004. URL: https://cds.cern.ch/record/782076 (cit. on pp. 1, 29, 31).

[4] ATLAS Collaboration. *The ATLAS Technical Design Report*. URL: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/TDR/access.html (cit. on pp. 1, 34, 36).

[5] A. N. Safonov. "CMS experiment at the LHC: Commissioning and early physics". In: *21st Rencontres de Blois on Windows on the Universe Blois, France, June 21-27, 2009*. 2010. arXiv: 1003.4038 [hep-ex]. URL: http://inspirehep.net/record/849640/files/arXiv:1003.4038.pdf (cit. on pp. 1, 35).

[6] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2013), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex] (cit. on pp. 1, 76).

[7] Serguei Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2013), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex] (cit. on p. 1).

[8] K. A. Olive et al. "Review of Particle Physics". In: *Chin. Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001 (cit. on pp. 1, 3, 7, 10, 19).

[9] Carlo Rovelli. "Notes for a brief history of quantum gravity". In: *Recent developments in theoretical and experimental general relativity, gravitation and relativistic field theories. Proceedings, 9th Marcel Grossmann Meeting, MG'9, Rome, Italy, July 2-8, 2000. Pts. A-C*. 2000, pp. 742–768. arXiv: gr-qc/0006061 [gr-qc]. URL: http://alice.cern.ch/format/showfull?sysnb=2193246 (cit. on pp. 3, 5).

[10] K Nakamura and Particle Data Group. "Review of Particle Physics". In: *Journal of Physics G: Nuclear and Particle Physics* 37.7A (2010), p. 075021. URL: http://stacks.iop.org/0954-3899/37/i=7A/a=075021 (cit. on p. 4).

[11] H. Nishino et al. "Search for Proton Decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in a Large Water Cherenkov Detector". In: *Phys. Rev. Lett.* 102 (14 Apr. 2009), p. 141801. DOI: 10.1103/PhysRevLett.102.141801. URL: http://link.aps.org/doi/10.1103/PhysRevLett.102.141801 (cit. on p. 5).

[12] S. K. Choi et al. "Observation of a resonance-like structure in the $\pi^\pm \psi'$ mass distribution in exclusive $B \to K \pi^\pm \psi'$ decays". In: *Phys. Rev. Lett.* 100 (2008), p. 142001. DOI: 10.1103/PhysRevLett.100.142001. arXiv: 0708.1790 [hep-ex] (cit. on p. 5).

[13] Roel Aaij et al. "Observation of the resonant character of the $Z(4430)^-$ state". In: *Phys. Rev. Lett.* 112.22 (2014), p. 222002. DOI: 10.1103/PhysRevLett.112.222002. arXiv: 1404.1903 [hep-ex] (cit. on p. 5).

[14] R. Aaij et al. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". In: *Phys. Rev. Lett.* 115 (7 Aug. 2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. URL: http://link.aps.org/doi/10.1103/PhysRevLett.115.072001 (cit. on p. 5).

[15] A Zee. *Quantum Field Theory in a Nutshell*. Nutshell handbook. Princeton, NJ: Princeton Univ. Press, 2003. URL: https://cds.cern.ch/record/706825 (cit. on p. 6).

[16] G. Drexlin et al. "Current direct neutrino mass experiments". In: *Adv. High Energy Phys.* 2013 (2013), p. 293986. DOI: 10.1155/2013/293986. arXiv: 1307.0101 [physics.ins-det] (cit. on p. 6).

[17] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.321 (cit. on p. 9).

[18] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.508 (cit. on p. 9).

[19] J. Goldstone. "Field Theories with Superconductor Solutions". In: *Nuovo Cim.* 19 (1961), pp. 154–164. DOI: 10.1007/BF02812722 (cit. on p. 9).

[20] G. Bernardi, M. Carena, and T. Junk. "Higgs Bosons: Theory and Searches". In: *Reviews of Particle Data Group: Hypothetical particles and Concepts* (2007) (cit. on pp. 10, 11).

[21] Tilman Plehn and Michael Rauch. "The Quartic Higgs Coupling at Hadron Colliders". In: *Phys. Rev.* D72 (2005), p. 053008. DOI: 10.1103/PhysRevD.72.053008. arXiv: hep-ph/0507321 [hep-ph] (cit. on p. 10).

[22] Robert Shrock. "Some recent results on models of dynamical electroweak symmetry breaking". In: *The origin of mass and strong coupling gauge theories. Proceedings, 5th International Workshop, SCGT'06, Nagoya, Japan November 21-24, 2006*. 2007, pp. 227–241. arXiv: hep-ph/0703050 [HEP-PH] (cit. on p. 10).

[23] Adam Martin. "Predicted Signals at the LHC from Technicolor: Erice Lecture". In: *Subnucl. Ser.* 46 (2011), pp. 135–159. DOI: 10.1142/9789814340212_0004. arXiv: 0812.1841 [hep-ph] (cit. on p. 10).

[24] Christopher T. Hill and Elizabeth H. Simmons. "Strong dynamics and electroweak symmetry breaking". In: *Phys. Rept.* 381 (2003). [Erratum: Phys. Rept.390,553(2004)], pp. 235–402. DOI: 10.1016/S0370-1573(03)00140-6. arXiv: hep-ph/0203079 [hep-ph] (cit. on p. 10).

[25] L. Álvarez-Gaumé and J. Ellis. "Eyes on a prize particle". In: *Nature Physics* 7 (Jan. 2011), pp. 2–3. DOI: 10.1038/nphys1874 (cit. on p. 12).

[26] Stephen P. Martin. "A Supersymmetry primer". In: *Adv. Ser. Direct. High Energy Phys.18,1(1998)* (1997). DOI: 10.1142/9789812839657_0001, 10.1142/9789814307505_0001. arXiv: hep-ph/9709356 [hep-ph] (cit. on pp. 14, 15).

[27] Georges Aad et al. "Search for top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 8\,\text{TeV}$ pp collisions with the ATLAS detector". In: *JHEP* 11 (2014), p. 118. DOI: 10.1007/JHEP11(2014)118. arXiv: 1407.0583 [hep-ex] (cit. on p. 14).

[28] Leszek Roszkowski, Enrico Maria Sessolo, and Andrew J. Williams. "What next for the CMSSM and the NUHM: Improved prospects for superpartner and dark matter detection". In: *JHEP* 08 (2014), p. 067. DOI: 10.1007/JHEP08(2014)067. arXiv: 1405.4289 [hep-ph] (cit. on p. 15).

[29] J. Alcaraz. *Precision Electroweak Measurements and Constraints on the Standard Model*. 2009. arXiv: 0911.2604 [hep-ex] (cit. on p. 16).

[30] TEVNPH Working Group. "Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb^{-1} of Data". In: *Proceedings, 46th Rencontres de Moriond on Electroweak Interactions and Unified Theories*. 2011. arXiv: 1107.5518 [hep-ex]. URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-11-354 (cit. on pp. 16, 27).

[31] Domènec Espriu and Brian Yencho. "Longitudinal WW scattering in light of the Higgs boson discovery". In: *Phys. Rev.* D87.5 (2013), p. 055017. DOI: 10.1103/PhysRevD.87.055017. arXiv: 1212.4158 [hep-ph] (cit. on p. 16).

[32] J. M. Blairon et al. "Chiral Symmetry Breaking in the Action Formulation of Lattice Gauge Theory". In: *Nucl. Phys.* B180 (1981), p. 439. DOI: 10.1016/0550-3213(81)90061-4 (cit. on p. 16).

[33] Urs M. Heller. "Status of the Higgs mass bound". In: *Nucl. Phys. Proc. Suppl.* 34 (1994), pp. 101–106. DOI: 10.1016/0920-5632(94)90323-9. arXiv: hep-lat/9311058 [hep-lat] (cit. on p. 16).

[34] LHC Higgs Cross-Section Working Group. *Standard Model Higgs boson production branching ratios*. URL: https://twiki.cern.ch/twiki/pub/LHCPhysics/LHCHXSWGCrossSectionsFigures/Higgs_BR_LM.png (cit. on p. 17).

[35] LHC Higgs Cross-Section Working Group. *Standard Model Higgs boson production cross sections*. URL: https://twiki.cern.ch/twiki/pub/LHCPhysics/CrossSectionsFigures/Higgs_XS_8TeV_lx.pdf (cit. on p. 18).

[36] Risto Raitio and Walter W. Wada. "Higgs-boson production at large transverse momentum in quantum chromodynamics". In: *Phys. Rev. D* 19 (3 Feb. 1979), pp. 941–944. DOI: 10.1103/PhysRevD.19.941. URL: http://link.aps.org/doi/10.1103/PhysRevD.19.941 (cit. on pp. 18, 19).

[37] S. Dittmaier et al. *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*. 2011. arXiv: 1101.0593 [hep-ph] (cit. on pp. 19, 20, 23, 93, 100).

[38] W. Beenakker et al. "Higgs Radiation Off Top Quarks at the Tevatron and the LHC". In: *Phys. Rev. Lett.* 87 (20 Oct. 2001), p. 201805. DOI: 10.1103/PhysRevLett.87.201805. URL: http://link.aps.org/doi/10.1103/PhysRevLett.87.201805 (cit. on p. 19).

[39] E. Richter-Was and M. Sapinski. "Search for the SM and MSSM Higgs boson in the $t\bar{t}H$ $H \to b\bar{b}$ channel". In: *ATL-PHYS-98-132* (1998) (cit. on p. 21).

[40] J. Cammin and M. Schumacher. "The ATLAS discovery potential for the channel $t\bar{t}H$, $H \to b\bar{b}$". In: *ATL-PHYS-2003-024* (2003) (cit. on p. 21).

[41] D. J. Miller et al. "Detecting heavy charged Higgs bosons at the CERN LHC with four b quark tags". In: *Phys. Rev.* D61 (2000), p. 055011. DOI: 10.1103/PhysRevD.61.055011. arXiv: hep-ph/9906230 [hep-ph] (cit. on p. 22).

[42] N Andari et al. *Higgs Production Cross Sections and Decay Branching Ratios*. Tech. rep. ATL-PHYS-INT-2010-030. Geneva: CERN, Mar. 2010. URL: https://cds.cern.ch/record/1255653 (cit. on p. 23).

[43] *LEP design report*. Geneva: CERN, 1984. URL: http://cds.cern.ch/record/102083 (cit. on p. 23).

[44] Stefan Roth. *Precision Electroweak Physics at Electron-Positron Colliders*. Springer tracts in modern physics. Berlin, Heidelberg: Springer, 2007. URL: https://cds.cern.ch/record/1338627 (cit. on p. 24).

[45] S. Schael et al. "Precision electroweak measurements on the Z resonance". In: *Phys. Rept.* 427 (2006), pp. 257–454. DOI: 10.1016/j.physrep.2005.12.006. arXiv: hep-ex/0509008 [hep-ex] (cit. on p. 24).

[46] ALEPH Collaboration et al. "Precision Electroweak Measurements and Constraints on the Standard Model". In: *ArXiv e-prints* (Dec. 2010). arXiv: 1012.2367 [hep-ex] (cit. on p. 24).

[47] R. Barate et al. "Search for the standard model Higgs boson at LEP". In: *Phys. Lett.* B565 (2003), pp. 61–75. DOI: 10.1016/S0370-2693(03)00614-2. arXiv: hep-ex/0306033 [hep-ex] (cit. on pp. 24, 25, 26).

[48] P. Teixeira-Dias (on behalf of the LEP Higgs working group). "Higgs boson searches at LEP". In: *J. Phys.: Conf. Ser. 110 042030* (2008) (cit. on p. 25).

[49] Qi-zhong Li. "Higgs searches at the Tevatron". In: *Proceedings, 38th Rencontres de Moriond on QCD and High-Energy Hadronic Interactions*. 2003. arXiv: hep-ex/0306006 [hep-ex]. URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-03-154 (cit. on pp. 25, 26).

[50] Georges Aad et al. "Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC". In: *Phys. Lett.* B726 (2013). [Erratum: Phys. Lett.B734,406(2014)], pp. 88–119. DOI: 10.1016/j.physletb.2014.05.011,10.1016/j.physletb.2013.08.010. arXiv: 1307.1427 [hep-ex] (cit. on p. 27).

[51] Georges Aad et al. "Evidence for the spin-0 nature of the Higgs boson using ATLAS data". In: *Phys. Lett.* B726 (2013), pp. 120–144. DOI: 10.1016/j.physletb.2013.08.026. arXiv: 1307.1432 [hep-ex] (cit. on p. 27).

[52] *Combination of standard model Higgs boson searches and measurements of the properties of the new boson with a mass near 125 GeV*. Tech. rep. CMS-PAS-HIG-13-005. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1542387 (cit. on p. 27).

[53] *Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector*. Tech. rep. ATLAS-CONF-2013-108. Geneva: CERN, Nov. 2013. URL: https://cds.cern.ch/record/1632191 (cit. on p. 27).

[54] The ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012) (cit. on p. 27).

[55] STFC. *CERN Accelerator Complex*. URL: http://www.stfc.ac.uk/research/particle-physics-and-particle-astrophysics/particle-physics/large-hadron-collider/cern-accelerator-complex/ (cit. on p. 30).

[56] C. M. Carloni Calame et al. "Large angle Bhabha scattering and luminosity at flavor factories". In: *Nucl. Phys.* B584 (2000), pp. 459–479. DOI: 10.1016/S0550-3213(00)00356-4. arXiv: hep-ph/0003268 [hep-ph] (cit. on p. 31).

[57] P. Hopchev. *LHCb Beam-Gas Imaging Results*. 2011. arXiv: 1107.1492 [hep-ex] (cit. on p. 31).

[58] S. M. White et al. *First Luminosity Scans in the LHC*. 2010. URL: http://accelconf.web.cern.ch/AccelConf/IPAC10/papers/mopec014.pdf (cit. on pp. 31, 32).

[59] Vladislav Balagura. "Notes on van der Meer Scan for Absolute Luminosity Measurement". In: *Nucl. Instrum. Meth.* A654 (2011), pp. 634–638. DOI: 10.1016/j.nima.2011.06.007. arXiv: 1103.1129 [physics.ins-det] (cit. on p. 31).

[60] Andre Sopczak et al. "MPX Detectors as LHC Luminosity Monitor". In: *IEEE Trans. Nucl. Sci.* 62 (2015), p. 3225. DOI: 10.1109/TNS.2015.2496316. arXiv: 1512.08014 [physics.ins-det] (cit. on p. 31).

[61] Nick Ryckx. "Combined energy ramp and beta star squeeze at the Large Hadron Collider". MA thesis. Ecole Polytechnique, Lausanne, 2012. URL: http://inspirehep.net/record/1186145/files/CERN-THESIS-2012-004.pdf (cit. on p. 32).

[62] Z Marshall. *Simulation of Pile-up in the ATLAS Experiment*. Oct. 2013. URL: https://cds.cern.ch/record/1605834 (cit. on pp. 32, 87).

[63] P. Lebrun. "Large Cryogenic Helium Refrigeration System for the LHC". In: (2003). URL: http://cds.cern.ch/record/605468/files/lhc-project-report-629.pdf (cit. on p. 33).

[64] P Cortese et al. *ALICE physics performance: Technical Design Report*. Technical Design Report ALICE. revised version submitted on 2006-05-29 15:15:40. Geneva: CERN, 2005. URL: https://cds.cern.ch/record/879894 (cit. on p. 34).

[65] K Aamodt et al. "First proton-proton collisions at the LHC as observed with the ALICE detector: Measurement of the charged particle pseudorapidity density at $\sqrt{s} = 900$ GeV". In: *Eur. Phys. J.* C65 (2010), pp. 111–125. DOI: 10.1140/epjc/s10052-009-1227-4. arXiv: 0911.5430 [hep-ex] (cit. on pp. 34, 37).

[66] G L Bayatian et al. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. There is an error on cover due to a technical problem for some items. Geneva: CERN, 2006. URL: http://cds.cern.ch/record/922757 (cit. on p. 35).

[67] G L Bayatian et al. "CMS Physics: Technical Design Report Volume 2: Physics Performance". In: *J. Phys. G* 34.CERN-LHCC-2006-021. CMS-TDR-8-2 (2007). revised version submitted on 2006-09-22 17:44:47, 995–1579. 669 p. URL: http://cds.cern.ch/record/942733 (cit. on p. 35).

[68] The LHCb Collaboration. *LHCb : Technical Proposal*. Tech. Proposal. Geneva: CERN, 1998. URL: https://cds.cern.ch/record/622031 (cit. on p. 35).

[69] ATLAS Collaboration. URL: http://www.atlas.ch/photos/full-detector-cgi.html (cit. on p. 39).

[70] The ATLAS Collaboration. *Overview of the ATLAS Inner Detector, with labels and dimensions.* URL: `http://www.atlas.ch/photos/inner-detector-combined.html` (cit. on p. 40).

[71] G Aad et al. *Expected performance of the ATLAS experiment: detector, trigger and physics.* Geneva: CERN, 2009. URL: `http://cds.cern.ch/record/1125884` (cit. on p. 42).

[72] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation 3* S08003.08 (2008) (cit. on p. 43).

[73] S. Eidelman. "Passage of Particles Through Matter". In: *Physics Letters B 592: 1* (2004) (cit. on p. 43).

[74] M. L. Andrieux et al. "Construction and test of the first two sectors of the ATLAS barrel liquid argon presampler". In: *Nuclear Instruments and Methods in Physics Research A* 479 (2002), pp. 316–333 (cit. on p. 44).

[75] P. Puzo. "ATLAS Calorimetry". In: *Nuclear Instruments & Methods in Physics Research A 494* (2002) (cit. on pp. 44, 45).

[76] M. J. Woudstra. "Precision of the ATLAS muon spectrometer". PhD thesis. University of Amsterdam, 2002 (cit. on p. 46).

[77] C. Adorisio et al. "System Test of the ATLAS Muon Spectrometer in the H8 Beam at the CERN SPS". In: *ATL-MUON-PUB-2007-005* (2007) (cit. on p. 46).

[78] C. Adorisio et al. "Study of the ATLAS MDT Spectrometer using High Energy CERN combined Test beam Data". In: *ATL-MUON-PUB-2008-005* (2008) (cit. on p. 46).

[79] J. Snuverink. "The ATLAS Muon Spectrometer: Commissioning and Tracking". PhD thesis. University of Twente, 2009 (cit. on p. 46).

[80] A. Buckley et al. "Rivet User Manual". Version 2.2.0. In: *Comput.Phys.Commun.* 184 (2013), pp. 2803–2819. URL: `http://rivet.hepforge.org/rivet-manual.pdf` (cit. on pp. 50, 52).

[81] Fabio Maltoni and Tim Stelzer. "MadEvent: Automatic event generation with MadGraph". In: *JHEP* 0302 (2003), p. 027 (cit. on p. 51).

[82] M.L. Mangano et. al. "Alpgen, a generator for hard multiparton processes in hadronic collisions". In: *JHEP* 0307 (2003), p. 001 (cit. on p. 51).

[83] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput.Phys.Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph] (cit. on p. 52).

[84] *Athena, The ATLAS Common Framework*. 2006. URL: http://atlas-computing.web.cern.ch/atlas-computing/documentation/swDoc/AthenaDeveloperGuide-8.0.0-draft.pdf (cit. on p. 52).

[85] *Gaudi Users Guide*. 2001. URL: http://lhcb-comp.web.cern.ch/lhcb-comp/Frameworks/Gaudi/Gaudi_v9/GUG/GUG.pdf (cit. on p. 52).

[86] G.P. Salam M. Cacciari and G. Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063 (cit. on pp. 53, 86).

[87] J. Allwall et al. *MadEvent – a multi-purpose minimal event generator powered by MadGraph – Minimal User Guide*. Tech. rep. 2007. URL: https://cp3.irmp.ucl.ac.be/projects/madgraph/attachment/wiki/ManualAndHelp/Manual-March-2007.pdf (cit. on p. 60).

[88] A. V. Belitsky and A. V. Radyushkin. "Unraveling hadron structure with generalized parton distributions". In: *Phys. Rept.* 418 (2005), pp. 1–387. DOI: 10.1016/j.physrep.2005.06.002. arXiv: hep-ph/0504030 [hep-ph] (cit. on p. 64).

[89] Pavel M. Nadolsky et al. "Implications of CTEQ global analysis for collider observables". In: *Phys. Rev.* D78 (2008), p. 013004. DOI: 10.1103/PhysRevD.78.013004. arXiv: 0802.0007 [hep-ph] (cit. on pp. 64, 91).

[90] ATLAS Collaboration. "ATLAS Tunes of PYTHIA6 and PYTHIA8 for MC11". In: *ATL-PHYS-PUB-2011-008* (2011) (cit. on pp. 64, 93).

[91] ATLAS Collaboration. "New ATLAS event generator tunes to 2010 data". In: *ATL-PHYS-PUB-2011-009* (2011) (cit. on pp. 64, 93).

[92] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph] (cit. on pp. 64, 91).

[93] Georges Aad et al. "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C75.7 (2015), p. 349. DOI: 10.1140/epjc/s10052-015-3543-1. arXiv: 1503.05066 [hep-ex] (cit. on pp. 64, 86, 87, 109, 121, 135, 140).

[94] J.M. Bernardo and A.F.M. Smith. *Bayesian Theory*. Wiley, 2000 (cit. on p. 67).

[95] Particle Data Group (G. Cowan). *The PDG review of statistical methods*. 2010. URL: http://pdg.lbl.gov/2010/reviews/rpp2010-rev-statistics.pdf (cit. on p. 67).

[96] L. Lyons. *Statistical Issues in Particle Physics*. 2008. URL: http://www.physics.ox.ac.uk/Users/lyons/R_H.pdf (cit. on p. 67).

[97] S. I. Bityukov and N. V. Krasnikov. *Towards the observation of signal over background in future experiments*. 1998. arXiv: physics/9808016 [physics] (cit. on p. 68).

[98] Pekka K. Sinervo. "Signal significance in particle physics". In: *Advanced Statistical Techniques in Particle Physics. Proceedings, Conference, Durham, UK, March 18-22, 2002*. 2002, pp. 64–76. arXiv: hep-ex/0208005 [hep-ex]. URL: http://www.ippp.dur.ac.uk/Workshops/02/statistics/proceedings//sinervo.pdf (cit. on p. 68).

[99] T. Junk. *Confidence Level Computation for Combining Searches with Small Statistics*. 1999. URL: http://root.cern.ch/root/doc/TomJunk.pdf (cit. on p. 68).

[100] E. Gross and O. Vitells. "Trial factors for the look elsewhere effect in high energy physics". In: *Eur.Phys.J.C70:525-530,2010* (2010) (cit. on p. 70).

[101] G. A. Young and R. L. Smith. *Essentials of Statistical Inference*. Cambridge University Press, 2005 (cit. on p. 72).

[102] D. Basu. "On the Elimination of Nuisance Parameters". In: *Journal of the American Statistical Association* (1977) (cit. on p. 72).

[103] N. Reid. "Aspects of modified profile likelihood". In: *Nonparametric Statistics and Related Topics, Elsevier* (1992) (cit. on p. 72).

[104] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *European Physical Journal C* 71, 1554 (Feb. 2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an] (cit. on pp. 72, 74, 83).

[105] J. Neyman and E. S. Pearson. "On the Problem of the most Efficient Tests of Statistical Hypotheses". In: *Physical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (1933) (cit. on p. 73).

[106] L Demortier and L Lyons. *Everything you always wanted to know about pulls*. Tech. rep. CDF/ANAL/PUBLIC/5776. CDF, Feb. 2002 (cit. on p. 74).

[107] Nicolas Chanon. *Multivariate analysis in high energy physics*. 2012. URL: https://people.phys.ethz.ch/~pheno/Lectures2012_StatisticalTools/slides/Chanon2.pdf (cit. on pp. 76, 79, 80).

[108] Byron P. Roe et al. "Boosted decision trees, an alternative to artificial neural networks". In: *Nucl. Instrum. Meth.* A543.2-3 (2005), pp. 577–584. DOI: 10.1016/j.nima.2004.12.018. arXiv: physics/0408124 [physics] (cit. on p. 77).

[109] G. Cowan. "Topics in statistical data analysis for high-energy physics". In: *High-energy physics. Proceedings, 17th European School, ESHEP 2009, Bautzen, Germany, June 14-27, 2009*. [,207(2013)]. 2013, pp. 197–218. DOI: 10.5170/CERN-2013-003.207,10.5170/CERN-2010-002.197. arXiv: 1012.3589 [physics.data-an]. URL: https://inspirehep.net/record/881668/files/arXiv:1012.3589.pdf (cit. on p. 77).

[110] R. Vilalta et al. "Automatic Signal Enhancement in Particle Physics using Multivariate Classification and Physical Constraints." In: *Ninth Workshop on Mining Scientific and Engineering Datasets (MSD06)* (2006). URL: http://www3.cs.uh.edu/~vilalta/papers/2006/siam06.pdf (cit. on p. 77).

[111] Mehmet Özgür Sahin, Dirk Krücker, and Isabell-Alissandra Melzer-Pellmann. *Performance and optimization of support vector machines in high-energy physics classification problems*. 2016. arXiv: 1601.02809 [hep-ex] (cit. on p. 77).

[112] J. Ocariz. "Probability and Statistics for Particle Physicists". In: *Proceedings, 1st Asia-Europe-Pacific School of High-Energy Physics (AEPSHEP)*. [,253(2014)]. 2014, pp. 253–280. DOI: 10.5170/CERN-2014-001.253. arXiv: 1405.3402 [physics.data-an]. URL: https://inspirehep.net/record/1296115/files/arXiv:1405.3402.pdf (cit. on p. 77).

[113] Liliana Teodorescu. *Artificial neural networks in high-energy physics*. 2008. URL: http://cds.cern.ch/record/1100521 (cit. on p. 77).

[114] A. Hoecker et al. "TMVA - Toolkit for Multivariate Data Analysis". In: *ArXiv Physics e-prints* (Mar. 2007). eprint: physics/0703039 (cit. on pp. 77, 81, 108).

[115] Bernd A. Berg. "New algorithm to investigate neural networks". In: *Comput. Phys. Commun.* 98 (1996), pp. 35–44. DOI: 10.1016/0010-4655(96)00086-0. arXiv: hep-ph/9604207 [hep-ph] (cit. on p. 78).

[116] G. Cybenko. "Approximations by superpositions of sigmoidal functions". In: *Mathematics of Control, Signals, and Systems* (1989) (cit. on p. 78).

[117] B.D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, 1996 (cit. on p. 79).

[118] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995 (cit. on pp. 80, 81).

[119] R. Doursat S. Geman E. Bienenstock. "Neural networks and the bias/variance dilemma". In: *Neural Computation 4: 1-58* (1992) (cit. on p. 80).

[120] Physics Information Technologies GmbH. *The NeuroBayes User's Guide*. 2010. URL: http://neurobayes.phi-t.de/nb_doc/NeuroBayes-HowTo.pdf (cit. on p. 82).

[121] W. Verkerke and D. Kirkby. *RooFit Users Manual v2.91*. URL: ftp://root.cern.ch/root/doc/RooFit_Users_Manual_2.91-33.pdf (cit. on p. 82).

[122] Aaron Armbruster. *runAsymptoticsCLs.C*. 2012. URL: http://rompotis.web.cern.ch/rompotis/atlas/higgs/limit/combination/CombPlots/Code/runAsymptoticsCLsFactory-00-01-02.cc (cit. on p. 83).

[123] F. James. *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*. Geneva, 1994. URL: https://inspirehep.net/record/1258343/ (cit. on p. 83).

[124] Georges Aad et al. "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data". In: *Eur. Phys. J.* C72 (2012), p. 1909. DOI: 10.1140/epjc/s10052-012-1909-1. arXiv: 1110.3174 [hep-ex] (cit. on p. 86).

[125] The ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (2008), S08003. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08003 (cit. on p. 86).

[126] C. et al Cojocaru. "Hadronic calibration of the ATLAS liquid argon end-cap calorimeter in the pseudorapidity region $1.6 < |\eta| < 1.8$ in beam tests". In: *Nuclear Instruments and Methods in Physics Research A* 531 (Oct. 2004), pp. 481–514. DOI: 10.1016/j.nima.2004.05.133. eprint: physics/0407009 (cit. on p. 86).

[127] Georges Aad et al. "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$TeV". In: *Eur. Phys. J.* C73.3 (2013), p. 2304. DOI: 10.1140/epjc/s10052-013-2304-2. arXiv: 1112.6426 [hep-ex] (cit. on pp. 86, 96, 97).

[128] *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*. Tech. rep. ATLAS-CONF-2011-102. Geneva: CERN, July 2011. URL: https://cds.cern.ch/record/1369219 (cit. on p. 87).

[129] V. Boisvert. "Top quarks objects definition and performance at ATLAS". In: *J. Phys. Conf. Ser.* 452.1 (2013), p. 012012. DOI: 10.1088/1742-6596/452/1/012012 (cit. on pp. 87, 96).

[130] K. Kondo. "Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. I. Method and Toy Models". In: *Journal of the Physical Society of Japan* 57 (1988), pp. 4126–4140. DOI: 10.1143/JPSJ.57.4126. URL: http://journals.jps.jp/doi/abs/10.1143/JPSJ.57.4126 (cit. on p. 88).

[131] M. Baak et al. "HistFitter software framework for statistical data analysis". In: *Eur. Phys. J.* C75 (2015), p. 153. DOI: 10.1140/epjc/s10052-015-3327-7. arXiv: 1410.1280 [hep-ex] (cit. on p. 90).

[132] Michelangelo L. Mangano et al. "ALPGEN, a generator for hard multiparton processes in hadronic collisions". In: *JHEP* 07 (2003), p. 001. DOI: 10.1088/1126-6708/2003/07/001. arXiv: hep-ph/0206293 [hep-ph] (cit. on p. 91).

[133] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph] (cit. on p. 91).

[134] G. Corcella et al. "HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)". In: *JHEP* 01 (2001), p. 010. DOI: 10.1088/1126-6708/2001/01/010. arXiv: hep-ph/0011363 [hep-ph] (cit. on p. 91).

[135] Francesco Caravaglios and Mauro Moretti. "An algorithm to compute Born scattering amplitudes without Feynman graphs". In: *Phys. Lett.* B358 (1995), pp. 332–338. DOI: 10.1016/0370-2693(95)00971-M. arXiv: hep-ph/9507237 [hep-ph] (cit. on p. 91).

[136] Michelangelo L. Mangano, Mauro Moretti, and Roberto Pittau. "Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b} + n$ jets as a case study". In: *Nucl. Phys.* B632 (2002), pp. 343–362. DOI: 10.1016/S0550-3213(02)00249-3. arXiv: hep-ph/0108069 [hep-ph] (cit. on p. 91).

[137] Georges Aad et al. "Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *JHEP* 07 (2013), p. 032. DOI: 10.1007/JHEP07(2013)032. arXiv: 1304.7098 [hep-ex] (cit. on p. 91).

[138] Kirill Melnikov and Frank Petriello. "Electroweak gauge boson production at hadron colliders through O(alpha(s)2)". In: *Phys. Rev.* D74 (2006), p. 114017. DOI: 10.1103/PhysRevD.74.114017. arXiv: hep-ph/0609070 [hep-ph] (cit. on p. 91).

[139] John M. Campbell and R. Keith Ellis. "An Update on vector boson pair production at hadron colliders". In: *Phys. Rev.* D60 (1999), p. 113006. DOI: 10.1103/PhysRevD.60.113006. arXiv: hep-ph/9905386 [hep-ph] (cit. on pp. 91, 100).

[140] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [hep-ph] (cit. on p. 91).

[141] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph] (cit. on p. 91).

[142] P. Nason S. Alioli and E. Re. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 040 (cit. on pp. 91, 92).

[143] Matteo Cacciari et al. "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation". In: *Phys. Lett.* B710 (2014), pp. 612–622. DOI: 10.1016/j.physletb.2012.03.013. arXiv: 1111.5869 [hep-ph] (cit. on pp. 91, 98).

[144] Peter Baernreuther, Michal Czakon, and Alexander Mitov. "Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$". In: *Phys. Rev. Lett.* 109 (2012), p. 132001. DOI: 10.1103/PhysRevLett.109.132001. arXiv: 1204.5201 [hep-ph] (cit. on pp. 91, 98).

[145] Michal Czakon and Alexander Mitov. "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels". In: *JHEP* 12 (2012), p. 054. DOI: 10.1007/JHEP12(2012)054. arXiv: 1207.0236 [hep-ph] (cit. on pp. 91, 98).

[146] Michal Czakon and Alexander Mitov. "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction". In: *JHEP* 01 (2013), p. 080. DOI: 10.1007/JHEP01(2013)080. arXiv: 1210.6832 [hep-ph] (cit. on p. 91).

[147] Michal Czakon, Paul Fiedler, and Alexander Mitov. "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$". In: *Phys. Rev. Lett.* 110 (2013), p. 252004. DOI: 10.1103/PhysRevLett.110.252004. arXiv: 1303.6254 [hep-ph] (cit. on p. 91).

[148] Michal Czakon and Alexander Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph] (cit. on p. 91).

[149] M. Botje et al. *The PDF4LHC Working Group Interim Recommendations.* 2011. arXiv: 1101.0538 [hep-ph] (cit. on p. 92).

[150] A. D. Martin et al. "Parton distributions for the LHC". In: *Eur. Phys. J.* C63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph] (cit. on p. 92).

[151] A. D. Martin et al. "Uncertainties on alpha(S) in global PDF analyses and implications for predicted hadronic cross sections". In: *Eur. Phys. J.* C64 (2009), pp. 653–680. DOI: 10.1140/epjc/s10052-009-1164-2. arXiv: 0905.3531 [hep-ph] (cit. on p. 92).

[152] Hung-Liang Lai et al. "New parton distributions for collider physics". In: *Phys. Rev.* D82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph] (cit. on p. 92).

[153] Jun Gao et al. "CT10 next-to-next-to-leading order global analysis of QCD". In: *Phys. Rev.* D89.3 (2014), p. 033009. DOI: 10.1103/PhysRevD.89.033009. arXiv: 1302.6246 [hep-ph] (cit. on p. 92).

[154] Richard D. Ball et al. "Parton distributions with LHC data". In: *Nucl. Phys.* B867 (2013), pp. 244–289. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph] (cit. on p. 92).

[155] N. Kidonakis. "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production". In: *Phys. Rev. D* 83 (2011), p. 091503 (cit. on p. 92).

[156] G. Bevilacqua et al. "HELAC-NLO". In: *Comput. Phys. Commun.* 184 (2013), p. 986 (cit. on p. 92).

[157] P. Golonka and Z. Was. "PHOTOS Monte Carlo: a precision tool for QCD corrections in Z and W decays". In: *Eur. Phys. J. C* 45 (2006), p. 97 (cit. on p. 93).

[158] S. Jadach. "TAUOLA – a library of Monte Carlo programs to simulate decays of polarised τ leptons". In: *Comput. Phys. Commun.* 64 (1991), p. 275 (cit. on p. 93).

[159] J. Forshaw J. Butterworth and M. Seymour. "Multiparton interactions in photoproduction at HERA". In: *Z. Phys. C* 72 (1996), p. 637 (cit. on p. 93).

[160] J. Allison et al. "Geant4 developments and applications". In: *IEEE Transactions on Nuclear Science* 53.1 (Feb. 2006), pp. 270–278. ISSN: 0018-9499. DOI: 10.1109/TNS.2006.869826 (cit. on p. 93).

[161] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on p. 93).

[162] Georges Aad et al. "Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC". In: *Eur. Phys. J.* C71 (2011), p. 1630. DOI: 10.1140/epjc/s10052-011-1630-5. arXiv: 1101.2185 [hep-ex] (cit. on p. 94).

[163] *In-situ jet energy scale and jet shape corrections for multiple interactions in the first ATLAS data at the LHC.* Tech. rep. ATLAS-CONF-2011-030. Geneva: CERN, Mar. 2011. URL: https://cds.cern.ch/record/1337780 (cit. on p. 96).

[164] *In-situ pseudorapidity intercalibration for evaluation of jet energy scale uncertainty using dijet events in proton-proton collisions at* $\sqrt{s} = 7$ TeV. Tech. rep. ATLAS-CONF-2011-014. Geneva: CERN, Mar. 2011. URL: http://cds.cern.ch/record/1334876 (cit. on p. 96).

[165] *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy* $\sqrt{s} = 7$ TeV. Tech. rep. ATLAS-CONF-2010-054. Geneva: CERN, July 2010. URL: https://cds.cern.ch/record/1281311 (cit. on p. 96).

[166] *In-situ pseudo-rapidity inter-calibration to evaluate jet energy scale uncertainty and calorimeter performance in the forward region.* Tech. rep. ATLAS-CONF-2010-055. Geneva: CERN, July 2010. URL: https://cds.cern.ch/record/1281312 (cit. on p. 96).

[167] T. Barillari. "Jet Energy Scale Uncertainties in ATLAS". In: CERN. Geneva: CERN, 2012. URL: http://cdsweb.cern.ch/record/1473875/files/ATL-PHYS-PROC-2012-137.pdf (cit. on p. 96).

[168] *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*. Tech. rep. ATLAS-CONF-2014-004. Geneva: CERN, Feb. 2014. URL: https://cds.cern.ch/record/1664335 (cit. on p. 97).

[169] *b-jet tagging calibration on c-jets containing D^{*+} mesons*. Tech. rep. ATLAS-CONF-2012-039. Geneva: CERN, Mar. 2012. URL: https://cds.cern.ch/record/1435193 (cit. on p. 97).

[170] *Measurement of the Mistag Rate with 5 fb^{-1} of Data Collected by the ATLAS Detector*. Tech. rep. ATLAS-CONF-2012-040. Geneva: CERN, Mar. 2012. URL: https://cds.cern.ch/record/1435194 (cit. on p. 97).

[171] *Measurement of top-quark pair differential cross-sections in the l+jets channel in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*. Tech. rep. ATLAS-CONF-2013-099. Geneva: CERN, Sept. 2013. URL: http://cds.cern.ch/record/1600778 (cit. on p. 98).

[172] Fabio Cascioli et al. "NLO matching for $t\bar{t}b\bar{b}$ production with massive b-quarks". In: *Phys. Lett.* B734 (2014), pp. 210–214. DOI: 10.1016/j.physletb.2014.05.040. arXiv: 1309.5912 [hep-ph] (cit. on p. 98).

[173] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph] (cit. on p. 98).

[174] Nikolaos Kidonakis. "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production". In: *Phys. Rev.* D83 (2011), p. 091503. DOI: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792 [hep-ph] (cit. on p. 100).

[175] Nikolaos Kidonakis. "NNLL resummation for s-channel single top quark production". In: *Phys. Rev.* D81 (2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028. arXiv: 1001.5034 [hep-ph] (cit. on p. 100).

[176] John M. Campbell and R. Keith Ellis. "$t\bar{t}W^{\pm}$ production and decay at NLO". In: *JHEP* 07 (2012), p. 052. DOI: 10.1007/JHEP07(2012)052. arXiv: 1204.5678 [hep-ph] (cit. on p. 100).

[177] M. V. Garzelli et al. "$t\bar{t}W^{\pm}$ and $t\bar{t}Z$ Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects". In: *JHEP* 11 (2012), p. 056. DOI: 10.1007/JHEP11(2012)056. arXiv: 1208.2665 [hep-ph] (cit. on p. 100).

[178] Catherine Bernaciak et al. "Fox-Wolfram Moments in Higgs Physics". In: *Phys. Rev.* D87 (2013), p. 073014. DOI: 10.1103/PhysRevD.87.073014. arXiv: 1212.4436 [hep-ph] (cit. on p. 102).

[179] Georges Aad et al. "Search for charged Higgs bosons in the $H^{\pm} \to tb$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *JHEP* 03 (2016), p. 127. DOI: 10.1007/JHEP03(2016)127. arXiv: 1512.03704 [hep-ex] (cit. on pp. 126, 133, 135).